



Restauration des pistes sonores optiques cinématographiques : approche par traitement d'images

Abdelâali Hassaïne

► To cite this version:

Abdelâali Hassaïne. Restauration des pistes sonores optiques cinématographiques : approche par traitement d'images. Mathématiques [math]. École Nationale Supérieure des Mines de Paris, 2009. Français. NNT : . pastel-00005981

HAL Id: pastel-00005981

<https://pastel.archives-ouvertes.fr/pastel-00005981>

Submitted on 12 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ED n° 431 : Information, communication, modélisation et simulation

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

T H È S E

pour obtenir le grade de
Docteur de l'École des Mines de Paris
Spécialité «Morphologie Mathématique»

Présentée et soutenue publiquement
par

Abdelâali HASSAÏNE

le 9 décembre 2009

**Restauration des Pistes Sonores Optiques Cinématographiques :
Approche par Traitement d'Images**

Directeurs de thèse : Bernard BESSERER et Etienne DECENCIÈRE

Jury :

Corinne VACHIER	<i>Rapporteur</i>
Rudolf GSCHWIND	<i>Rapporteur</i>
Bernard BESSERER	<i>Examineur</i>
Jean-Hugues CHENOT	<i>Examineur</i>
Etienne DECENCIÈRE	<i>Examineur</i>
Claire-Hélène DEMARTY	<i>Examineur</i>
Dominique JEULIN	<i>Examineur</i>
Fernand MEYER	<i>Examineur</i>

Remerciements

Je tiens à remercier Corrine Vachier, Rudolf Gschwind, Jean-Hugues Chenot, Claire-Hélène Demarty, Dominique Jeulin et Fernand Meyer d'avoir accepté de juger ce travail.

Je voudrais aussi remercier mon directeur de thèse Etienne Decencière de qui, j'ai énormément appris. Ses conseils instructifs et ses discussions constructives m'ont beaucoup aidé durant cette thèse. Je le remercie également pour ses encouragements ainsi que pour la confiance qu'il a su m'accorder et qui m'a permis de savourer le travail scientifique durant ces trois années de thèse.

J'aimerais également remercier Bernard Besserer, mon autre directeur de thèse, d'avoir pris le temps à chaque fois de me guider, de m'expliquer et de me conseiller. Je le remercie également pour m'avoir accueilli plusieurs fois à La Rochelle.

Je voudrais exprimer aussi mes sincères remerciements à Christian Comte et Nicolas Ricordel, partenaires du projet RESONANCES avec qui, j'ai pu échanger et qui m'ont fait bénéficier de leurs conseils et expertise concernant les pistes sonores optiques.

Un grand merci également à toutes les personnes qui ont accepté de prendre de leur temps et de me prêter leurs oreilles en participant à la campagne de tests d'écoute en aveugle.

Le Centre de Morphologie Mathématique offre un environnement de travail idéal et une ambiance très agréable, et ce, grâce à la sympathie et la disponibilité de tous ses membres, nouveaux et anciens, que je tiens ici à remercier.

J'adresse mes remerciements tout particuliers à Catherine Moysan, l'ange gardien du CMM. Sa gentillesse, sa sympathie et sa disponibilité m'ont été d'une grande aide durant ces trois années.

Je souhaiterais également remercier Fernand Meyer de m'avoir accueilli au sein de son équipe et avec qui j'ai pu avoir des échanges très fructueux.

Je remercie aussi tous les chercheurs du CMM avec qui j'ai pu échanger, notamment Beatriz Marcotegui, qui ne manque jamais d'idées originales, Serge Beucher, à qui je dois plusieurs éclaircissements sur les formes et les squelettes, Petr Dokladal qui m'a fait découvrir les ensembles de niveaux, Jesus Angulo qui m'a offert l'occasion de manipuler des images multispectrales ainsi que Guillaume Thibault pour ses nombreux conseils en apprentissage automatique et classification.

Je suis également très reconnaissant à Michel Bilobeau pour m'avoir aidé à régler plusieurs problèmes informatiques et effectuer plusieurs choix techniques.

Un grand merci également aux anciens du CMM pour leurs conseils et leurs aides

concernant l'utilisation de Morph-M, notamment Thomas Retornaz, Raffi Enficiaud, Thibault Nion et Romain Lerallut.

Je remercie également tous mes autres collègues du CMM : Jean-Claude, Matthieu, Romaric, François, Serge K., Erik, Estelle, Jean S., Guillaume, Christophe, Jorge, Adnan, Louise, Charles, Hellen, Patrick, Jean F., Bruno, Julie, Santiago, Luc, Vincent et Arturo. Merci à vous pour toutes ces pauses café, ces discussions et ces plaisanteries.

Mes sincères remerciements à Salim Bensmina, Bachir Bouali et mon père qui ont pris de leur temps pour la relecture de cette thèse.

Outre le côté professionnel, ces trois années bellifontaines ont été pour moi une expérience personnelle très enrichissante, et ce, grâce à plusieurs personnes.

Je suis content d'avoir fait la connaissance de Salim Mokraoui. Je le remercie, ainsi que toute sa famille, pour tous ces accueils chaleureux, ces discussions et ces plats toujours très réussis.

Je remercie mon ex-colocataire Waheed Afzal, grâce à qui, j'ai découvert beaucoup de la culture pakistanaise : le Beriani, les Pakoras, le Gajar ka Halwa et même Coco Corina.

J'ai pu également apprécier pendant cette période, la joie de vivre batnie de Hakim, la gentillesse constantinoise de Raouf, la sympathie algéroise des deux Salim et de Djamel, le p'tit accent fassi (heu... casaoui) et la générosité meknassi des deux Otmane. Sans oublier l'humour ardéchois de Xavier, l'originalité parisienne de Gilles, les plaisanteries essonniennes de Matthieu, le papotage néerlandais d'Erik et le Français très... original de Christophe.

Je n'oublierais pas de remercier tous ceux avec qui j'ai partagé ces bons moments de foot et ces parties de volley, même si je ne suis pas très doué dans ces sports.

Je remercie également ma tante et ma grand-mère pour leur visite qui m'a beaucoup fait plaisir, ainsi qu'Abderrahmane, Adila, Djawad et Faïza pour ces nombreuses invitations et ces accueils très généreux.

Enfin, je tiens à remercier mes amis, ainsi que toute ma famille et mes parents en particulier, pour m'avoir encouragé à aller au bout de ma volonté et soutenu moralement pendant toutes ces longues heures de conversations téléphoniques.

Abstract

Restoration of optical soundtracks in motion pictures : the digital image-processing approach

The restoration of old motion picture films is an active research field. The restoration of the optical soundtrack however has only been performed in the audio domain, in spite of the fact that it is recorded as a continuous image located between the images of the film and the sprocket holes.

The restoration of the optical soundtrack directly at the image level has several advantages. First of all, the defects are visible, therefore we do not need to know their audio equivalent to be able to restore them. Moreover, the optical soundtrack has several geometrical properties that can be used for its restoration. Finally, such a restoration makes it possible to recover the original version of the soundtrack, since it only cancels the defects which are due to the film ageing.

Two types of optical soundtracks exist : the most common are variable area soundtracks, which contain a transparent region located between two symmetrical opaque regions. The width of the transparent region is directly proportional, at each instant, to the audio signal amplitude. For the restoration of this kind of soundtracks, we have developed algorithms for the correction of azimuth deviation, the detection of the symmetry axis and for the symmetry enforcement as well. Then, a segmentation step is performed for defects removal, and finally, an anti-aliasing correction is applied in order to cancel the noise which is due to the segmentation step.

The second type is called variable density soundtracks, in which the intensity of each line is proportional, at each instant, to the audio signal amplitude. For the restoration of those soundtracks, we have developed an algorithm for the detection and the correction of azimuth deviation. Then, a clipping algorithm is performed for defects removal. This last algorithm is parameterized automatically in such a way that it minimizes the total variation of the resulting audio signal and thus minimizes the noise.

We have also studied the problem of badly exposed optical soundtracks. This problem is due to light propagation during film duplications. Its effects are present in both image and audio domains. We have developed several indicators at the image level for the detection of badly exposed optical soundtracks. These indicators are based on the morphological properties of this problem on one side, and on the dissymmetry which it

introduces to soundtrack edges on the other side. A comparison between these indicators has been performed in order to study the credibility of each one of them. Finally, two restoration methods have been developed. The first one is a morphological approach that makes it possible to restore the shapes of peaks and valleys of the soundtrack. For less damaged cases, the second method matches the gray levels of the image against a look up table corresponding to the detected amount of bad exposure.

We have also discussed how the parallelization and the communication are performed between all the developed modules and studied the problem of their integration into the whole restoration system.

The evaluation of our restoration algorithms has been carried out by sound experts. We have also set up blind listening tests in order to have a more objective assessment.

Keywords

Restoration of optical soundtracks, mathematical morphology, region growing, noise reduction, exposure correction, blind listening tests.

Résumé

La restauration des films anciens est un domaine de recherche qui a été largement étudié par la communauté scientifique. La restauration de la piste son optique, quant à elle, n'a été en pratique effectuée que dans le domaine audio, bien qu'elle constitue une image continue sur la pellicule photographique, située entre la succession des images et les perforations.

La restauration par traitement d'images de la piste son a plusieurs intérêts. Tout d'abord, les défauts sont visibles, on n'a donc pas forcément besoin de connaître leurs équivalents audio pour les restaurer. En plus, la piste son présente plusieurs propriétés géométriques pouvant servir à sa restauration. Enfin, un tel type de restauration permet un retour à la version originelle de la piste son, puisqu'il ne supprime que les défauts dus au vieillissement du film.

Deux types de pistes sonores optiques existent : les plus répandues sont les pistes son à largeur variable qui contiennent une région transparente située entre deux régions opaques symétriques. La taille de la région transparente est proportionnelle, en chaque instant, à l'amplitude du signal audio. Pour la restauration de ce type de pistes, nous avons proposé des algorithmes pour le réglage d'azimut et la détection de l'axe de symétrie, ainsi que pour le renforcement de la symétrie. Une étape de segmentation est ensuite effectuée pour supprimer les différents défauts. Enfin, une correction anti-aliasing est apportée pour annuler le bruit généré par l'étape de segmentation.

Le second type est celui des pistes son à densité variable, dans lesquelles l'intensité de chaque ligne est proportionnelle, en chaque instant, à l'amplitude du signal audio. Pour la restauration de ce type de pistes, nous avons proposé un algorithme pour la détection et la correction de la déviation d'azimut, puis un algorithme de «clipping» pour la suppression des différents défauts. Ce dernier algorithme est automatiquement paramétré de sorte à minimiser la variation totale du signal audio restauré et ainsi minimiser le bruit.

Nous avons également étudié le problème de la mauvaise exposition des pistes son. Ce problème est dû à la diffusion de la lumière pendant les différentes copies. Ses effets sont présents aussi bien dans le domaine image que dans le domaine audio. Plusieurs indicateurs au niveau image ont été développés pour la détection des pistes son mal exposées. Ces indicateurs sont basés, d'une part, sur les propriétés morphologiques du problème en question, et d'autre part, sur la dissymétrie qu'il introduit au niveau des contours de la piste son. Une comparaison entre ces indicateurs a été effectuée pour étudier la crédibilité de chacun d'eux. Enfin, deux méthodes de restauration ont été développées, la première étant une approche morphologique permettant de corriger la forme des pics et des vallées de la piste son. Pour les cas moins sévères, la deuxième méthode applique aux niveaux de gris de l'image une table de correspondance (LUT) selon le degré de mauvaise exposition

déecté.

Nous avons abordé également le problème de la parallélisation et de la communication entre les différents modules réalisés. Nous avons aussi étudié le problème de leur intégration dans un système global de restauration.

Pour évaluer notre système, nous avons d'abord demandé l'avis de nos partenaires experts en la matière. Nous avons ensuite organisé une campagne d'évaluation en utilisant des tests audio en aveugle dans le but d'avoir une évaluation objective des résultats.

Mots clés

Restauration des pistes son optiques, morphologie mathématique, croissance de régions, réduction de bruit, correction d'exposition, tests d'écoute en aveugle.

Table des matières

1	Introduction	1
1.1	Contexte	1
1.2	Bref historique du son au cinéma	2
1.3	La piste son optique	2
1.4	Dégradation des pistes son optiques	5
1.5	Restauration des pistes son optiques	5
1.6	Le projet RESONANCES	7
1.7	Tests d'écoute en aveugle	10
1.8	Structure de cette thèse	13
1.9	Conventions et notations	14
2	Restauration des pistes à largeur variable	15
2.1	Définitions	15
2.1.1	Propriétés des pistes à largeur variable	15
2.1.2	Dynamique audio	15
2.2	Travaux antérieurs	16
2.2.1	Brevets	16
2.2.2	Travaux académiques	17
2.2.3	Méthode d'E. Brun	18
2.2.4	Bilan de l'étude bibliographique	19
2.3	Méthode proposée	20
2.3.1	Détection de l'axe de symétrie	20
2.3.2	Correction de l'azimut	22
2.3.3	Renforcement de la symétrie	28
2.3.4	Segmentation	30
2.3.5	Renforcement de causalité	32

2.3.6	Lissage des contours	32
2.3.7	Récapitulatif	35
2.3.8	Restauration des autres types de pistes VA	36
2.4	Résultats	37
2.4.1	Dans le domaine image	37
2.4.2	Dans le domaine audio	37
2.4.3	Temps d'exécution	39
2.5	Conclusion	39
3	Restauration des pistes à densité variable	41
3.1	Définitions	41
3.1.1	Propriétés des pistes à densité variable	41
3.1.2	Dynamique audio	41
3.2	Travaux antérieurs	42
3.3	Bases de données	43
3.4	Méthode proposée	44
3.4.1	Correction d'azimut	45
3.4.2	Correction des autres défauts	46
3.5	Résultats	48
3.5.1	Dans le domaine image	48
3.5.2	Dans le domaine audio	50
3.6	Conclusion et perspectives	51
4	Etude des pistes son mal exposées	53
4.1	Définitions	53
4.1.1	Origines et effets de la mauvaise exposition	53
4.1.2	Test de cross-modulation	55
4.1.3	Simulation de la mauvaise exposition	56
4.2	Travaux antérieurs	56

4.3	Constitution de la base de données expérimentale	57
4.4	Détection de la mauvaise exposition	58
4.4.1	Invariance par rapport à l'ouverture (ou la fermeture) morphologique	58
4.4.2	Dissymétrie des pics et des vallées	61
4.4.3	Etude comparative des différents indicateurs	65
4.4.4	Stabilité des indicateurs par rapport au seuil	71
4.5	Correction de la mauvaise exposition	73
4.5.1	Application d'une érosion (ou d'une dilatation)	73
4.5.2	L'application d'une table de correspondance (LUT) sur les niveaux de gris	79
4.5.3	Analyse des deux méthodes	80
4.6	Tests d'écoute	82
4.7	Conclusion	83
5	Système de restauration global	85
5.1	Communication et ordonnancement	85
5.2	Temps de calcul détaillés	90
5.3	Implémentation et possibilités de parallélisation	93
5.4	Conclusion	94
6	Conclusion générale	95
6.1	Apports de cette thèse	95
6.2	Perspectives	96
	Bibliographie	97

Introduction

1.1 Contexte

On estime aujourd’hui que 70 à 80% des films produits avant le début des années trente sont perdus à tout jamais. Après cette période, la perte des films est devenue moins dramatique mais est restée toujours considérable. Même les négatifs originaux de plusieurs films relativement récents ont été perdus, ou sont dans une condition déplorable [52].

Les causes de cette situation sont multiples : certains films ont été détruits quand ils n’étaient plus rentables, d’autres films muets ont été détruits quand les films sonores ont été introduits. De même, quand les pellicules en nitrate inflammable sont devenues prohibées, les films enregistrés sur ce support ont également été détruits.

De plus, les films ont la propriété de s’autodétruire. En effet, les films en nitrate de cellulose, qui étaient utilisés jusqu’au début des années cinquante, sont chimiquement très instables. Pour cette raison, une copie de sauvegarde était faite sur un support en acétate de cellulose pendant plusieurs décennies. On sait actuellement que ce support se décompose également par un procédé appelé «syndrome du vinaigre» à cause de l’odeur qu’il produit.

Les films anciens représentent un grand patrimoine culturel qui intéresse aujourd’hui beaucoup de personnes. Plusieurs départements universitaires et associations sont spécialisés dans le cinéma et son histoire. De plus, plusieurs festivals présentent exclusivement des films anciens. Pour ces raisons, on devient de plus en plus conscient de l’urgence de restaurer les films avant qu’ils ne soient complètement détruits.

En conséquence, une grande importance a été accordée à la restauration des films par les archives du film ainsi que par la communauté scientifique (voir à titre d’exemple [23, 41]). Cette restauration concerne aussi bien le son que les images du film. Cependant, la restauration de la piste sonore optique est généralement effectuée dans le domaine audio, malgré le fait qu’elle soit enregistrée sur la pellicule comme une image continue entre la succession des images et les perforations.

1.2 Bref historique du son au cinéma

Le monde du cinéma a connu bien des inventions avant l'introduction de la piste sonore optique. Le premier précurseur du cinéma sonore est sans doute le «kinétophone» inventé par T. Edison en 1895 [6]. Cet appareil était composé d'un «kinétoscope» permettant la visualisation du film par une seule personne, et d'un cylindre phonographique permettant la lecture du son. Ce n'est que quelques années plus tard, pendant l'exposition universelle de 1900 à Paris, que le système «Phono-Cinéma-Théâtre» de Clément-Maurice et Henri Lioret a permis de réaliser les premières projections alliant son et image [2].

Depuis, plusieurs brevets sur des dispositifs de projection de films sonores se sont succédés, mais ces dispositifs avaient cependant des problèmes de synchronisation, puisque le son et l'image étaient enregistrés sur des supports différents. En 1927, Warner Bros sort «The Jazz Singer» le premier «film parlant» utilisant un «vitaphone» pour la synchronisation du son avec les images [21]. Ce dispositif utilisait une technologie dite «son-sur-disque», c'est-à-dire que le son était enregistré sur un phonographe relié au projecteur par un mécanisme d'interdiction permettant ainsi la synchronisation. Ce mécanisme nécessitait cependant une constante supervision, parce que tout saut de disque ou changement de vitesse de projection pouvait causer un problème de synchronisation.

La technologie «son sur la pellicule» s'est développée parallèlement à la technologie son-sur-disque et a fini par s'imposer, parce qu'elle posait moins de problèmes de synchronisation. Le son accompagnant l'image était physiquement enregistré sur la pellicule. Cet enregistrement pouvait se faire, soit de manière optique, soit de manière magnétique. Bien que le support magnétique produisait initialement un son de meilleure qualité, il a fini par être abandonné à cause de son coût important et de sa courte durée de vie. A la fin des années quatre-vingt, Dolby a apporté plusieurs améliorations à la piste son optique, pour loger deux pistes qui, non seulement contiennent les deux canaux du son stéréo, mais sont encodées de sorte à simuler un troisième canal central et un quatrième qui véhicule le son ambiant et les effets spéciaux. Cette technologie porte le nom de Dolby SR (pour Spectral Recording) et est universellement utilisée jusqu'au jour d'aujourd'hui [25].

En 1992, Dolby a introduit la piste son numérique en utilisant l'espace situé entre les perforations. Le Dolby Digital Surround EX est le dernier développement de cette technologie, il date de 1999 et est le système numérique le plus répandu actuellement [33]. Cependant, la piste analogique est toujours présente sur les pellicules pour qu'elle puisse être lue dans les cinémas ne disposant pas de projecteur numérique (fig. 1.1).

1.3 La piste son optique

On parle de piste son optique quand le son est enregistré sur la pellicule de manière optique. Les procédés les plus connus utilisant cette technologie sont :

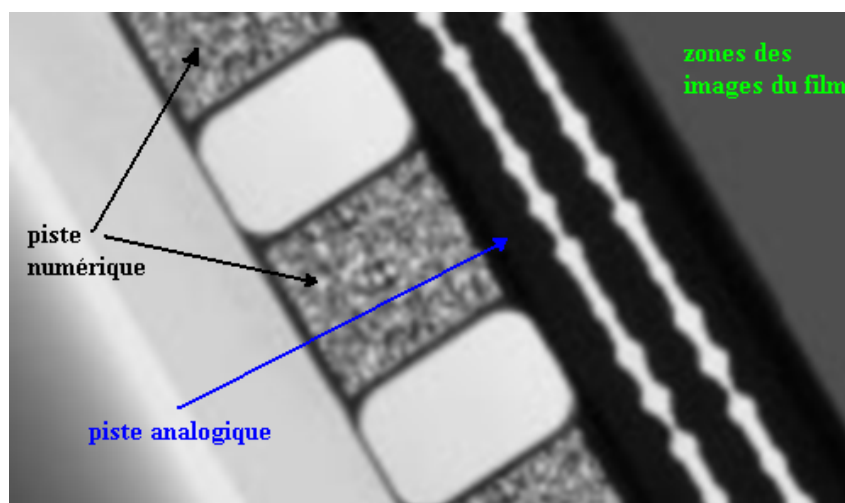


FIGURE 1.1 – Pistes analogique et numérique sur une pellicule cinématographique.

- les pistes à largeur variable (dites aussi à densité fixe, et connues anciennement sous le nom de «RCA Photophones» [19]). L'amplitude du signal audio dans ces pistes est modulée par une région transparente. Quand cette région est située à l'extrémité de la piste, on parle de pistes son à largeur variable unilatérales (fig. 1.2(a)). Quand elle est située entre deux régions opaques symétriques, on parle de pistes son à largeur variable bilatérales (fig. 1.2(b)). Selon le nombre de pistes son véhiculées, elles peuvent être mono optiques, stéréo optiques (fig. 1.2(c)) ou même multiples (fig. 1.2(d)) ;
- et les pistes son à densité variable (connues anciennement sous le nom de «movietones» [19]) qui contiennent des lignes parallèles avec un niveau de densité proportionnel à l'amplitude du signal audio. Ces pistes peuvent être aussi mono optiques (fig. 1.2(e)) ou stéréo optiques (fig. 1.2(f)).

Le procédé de lecture de la piste son optique est simple : pendant la projection du film, la piste son est éclairée par une lumière via une fente de lecture. La quantité de lumière qui passe à travers la piste son est convertie par une cellule photoélectrique en un courant électrique. Ce courant est ensuite converti en signal sonore à travers une chaîne d'amplification classique (fig. 1.3). Notons que ce même procédé permet de lire aussi bien les pistes à largeur variable que les pistes à densité variable [27].

Le son obtenu à partir des pistes à largeur variable est de qualité bien meilleure que celui qui correspond aux pistes à densité variable. Pour cette raison, les pistes à densité variable ont rapidement été abandonnées.

La standardisation de la reproduction du son a toujours été d'une grande nécessité : le son produit par différents studios doit être similaire dans toutes les salles de cinéma. Pour cette raison, le système sonore d'une salle de cinéma a été divisé en deux parties : la

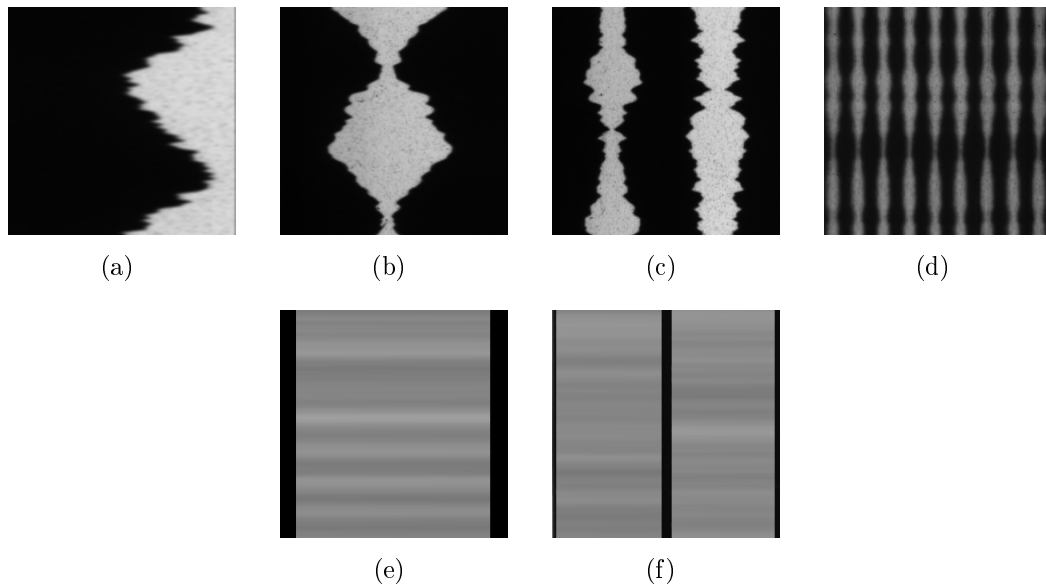


FIGURE 1.2 – Types de pistes son optiques.

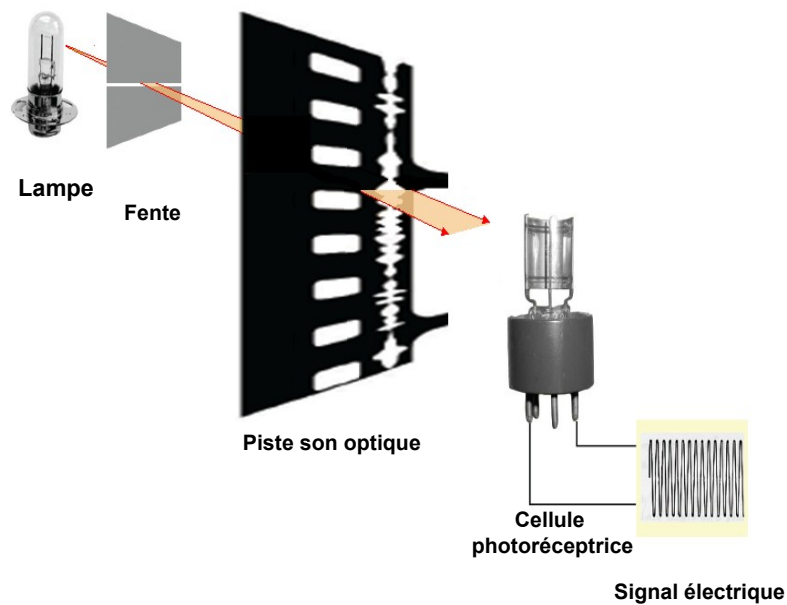


FIGURE 1.3 – Procédé de lecture d'une piste son optique.

«A-chain», qui concerne l'enregistrement et la reproduction du son [34], et la «B-chain», qui concerne les amplificateurs, les haut-parleurs et l'acoustique [35]. La réponse spectrale des pistes son spécifiée dans le norme A-chain s'appelle la X-Curve (pour eXtended range Curve). Cette courbe est plate jusqu'à 2 kHz, elle descend de 3 dB par octave jusqu'à 10 kHz, puis de 6 dB par octave, comme illustré dans la figure 1.4.

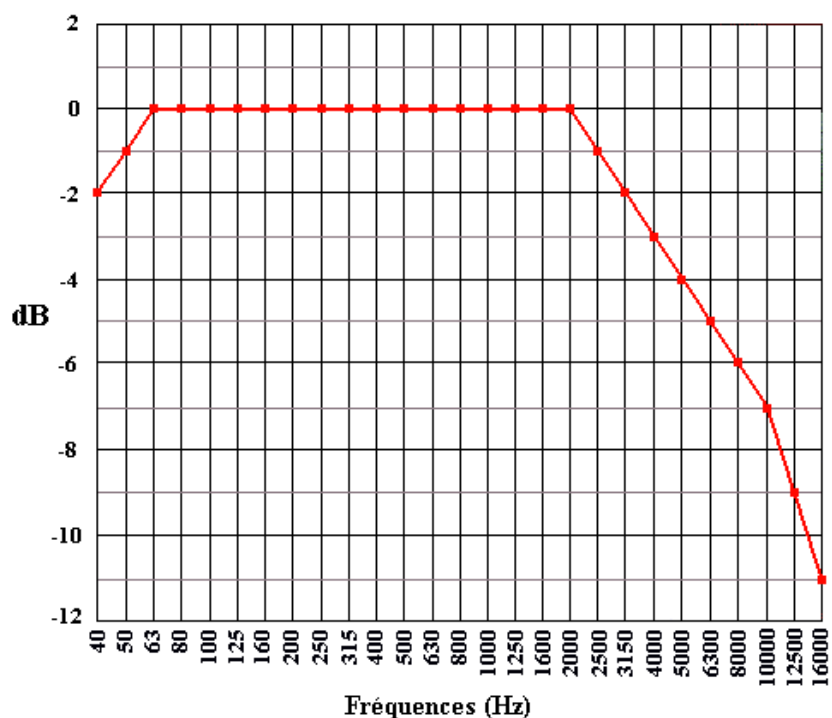


FIGURE 1.4 – Allure de la X-Curve.

1.4 Dégradation des pistes son optiques

La piste son optique étant enregistrée sur le même support que les images du film, elle subit, en conséquence, le même type de dégradations que l'on trouve sur une pellicule cinématographique : poussières, moisissures, taches, rayures ...etc. De plus, la piste son est quelquefois très dégradée, par abrasion du film (la piste son se trouve près des perforations). La figure 1.5 illustre quelques exemples de pistes son dégradées.

La mauvaise exposition des pistes son est un autre type de défauts, il est dû à la propagation de lumière durant les différentes copies. L'effet de la mauvaise exposition est plus clair sur les pistes à largeur variable. Remarquons, par exemple, que la piste sous-exposée de la figure 1.6 a des «pics» très pointus.

1.5 Restauration des pistes son optiques

Pour faire face aux dégradations des pistes son optiques, celles-ci sont généralement d'abord converties en signal audio. Ce signal est ensuite restauré par des méthodes numériques de traitement du signal [52].

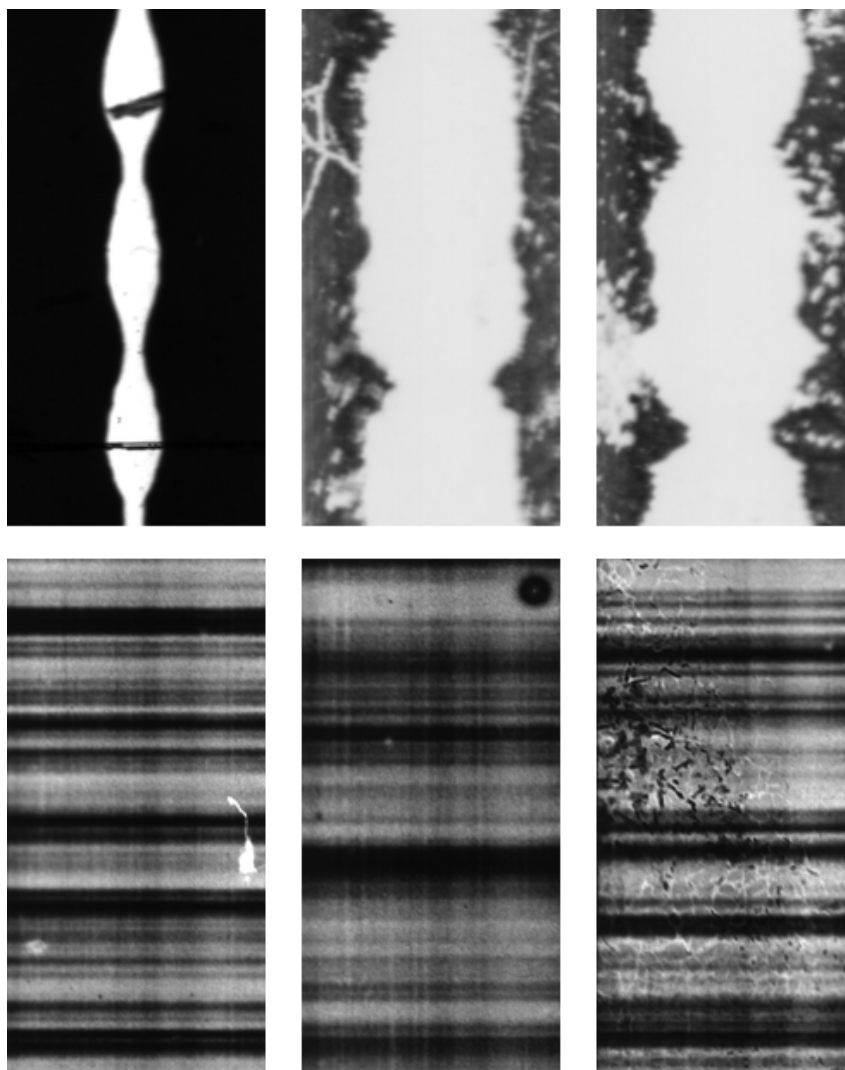


FIGURE 1.5 – Exemples de pistes son dégradées. *Les détails de la résolution et de l'échelle seront donnés ultérieurement.*

De nos jours, plusieurs entreprises et laboratoires de recherche travaillent sur des méthodes de traitement du signal qui ont bien apporté leurs fruits dans le domaine de la restauration audio ([28, 72]). Cependant, aussi performantes qu'elles soient, ces méthodes ne permettent pas de distinguer entre quelques artefacts audio dus à la dégradation du film, et d'autres sons originellement présents dans la piste son.

La restauration de la piste son optique directement dans le domaine image aurait, quant à elle, plusieurs avantages :

- Tout d'abord, les défauts sont visibles, il n'est donc pas nécessaire de connaître leur équivalents dans le domaine audio pour pouvoir les restaurer.
- De même, l'information géométrique peut être utilisée pour la restauration. Dans

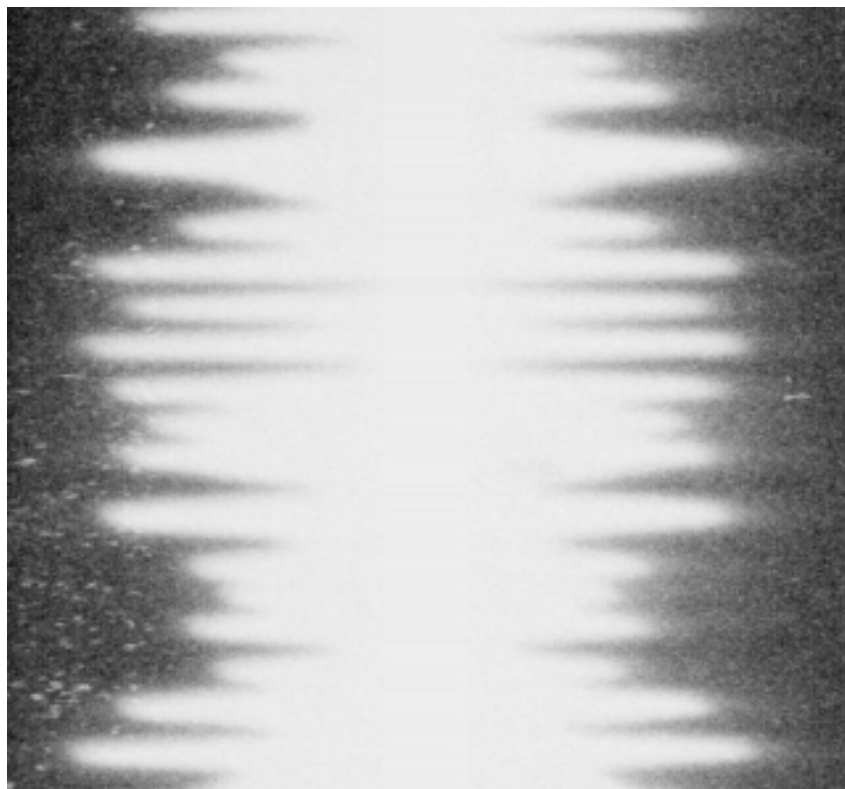


FIGURE 1.6 – Exemple de piste sous-exposée.

les pistes à largeur variable, par exemple, le fait que les deux côtés sont symétriques peut être utilisé pour restaurer le côté dégradé, en utilisant son côté symétrique. Quant aux pistes à densité variable, le fait que les lignes ont un niveau de gris constant peut aussi être utilisé pour localiser les défauts.

- Enfin, ce type de restauration permet un retour à la version originelle de la piste son, puisqu'il permet de distinguer entre les artefacts dus à la dégradation et les sons originellement présents. Ce dernier point est d'une importance cruciale pour les archivistes du film.

Il est important de noter que ce type de restauration n'est pas destiné à remplacer la restauration dans le domaine audio. Ces deux approches de restauration sont en effet complémentaires.

1.6 Le projet RESONANCES

C'est dans cet ordre d'idées que le projet RESONANCES (REstauration du SON par ANalyse de séquenCES d'images) a été lancé en 2006. Ce projet regroupe les quatre partenaires suivants :

- **Général de Travaux Cinématographiques (GTC)**, représenté par son directeur technique, Christian Comte. Ce laboratoire assure toute la chaîne de production et de copie des films sur les pellicules cinématographiques. Dans le cadre du projet RESONANCES, GTC fournit une partie du corpus, participe à l'intégration de la caméra linéaire, agit en tant que conseiller concernant les normes et usages des différentes pistes sonores optiques et contribue à la validation des résultats.
- **Centre National de Cinématographie/ Archives Françaises du Film(AFF)**, représenté par Nicolas Ricordel, responsable du pôle numérique du service laboratoire et restauration. Les AFF assurent l'ensemble des missions liées aux collections de films confiées à l'Etat. Dans le cadre de ce projet, les AFF ont les mêmes rôles que le GTC.
- **Université de la Rochelle/ Laboratoire Informatique, Image, Interaction (L3I)**, le L3I est impliqué depuis plusieurs années dans la restauration des films anciens. Il participe dans le cadre de ce projet à la recherche fondamentale, puis prend en charge le portage et l'intégration des outils, ainsi que le développement de la plate-forme d'acquisition avec caméra linéaire.
- **Ecole des Mines de Paris/ Centre de Morphologie Mathématique (CMM)**, le CMM a été impliqué dans la restauration des films anciens via le projet ESPRIT-NOBLESSE. Son rôle au sein du projet RESONANCES est le développement d'algorithmes automatiques pour la restauration des pistes optiques.

Défileur utilisé

La figure 1.7 montre le défileur utilisé dans le cadre de ce projet. Ce défileur a été créé en adaptant un «sepmag player» (appareil permettant la lecture des pistes magnétiques séparées). La capture est assurée par une caméra linéaire utilisant une optique macro de Schneider-Kreuznach. Cette caméra permet la capture de 48000 lignes de 512 pixels par seconde avec un niveau de gris codé sur 12 bits. L'éclairage est, quant à lui, assuré par une diode électroluminescente (LED).

Caractéristiques

Actuellement, les enregistreurs cinématographiques bénéficient d'une bande passante de 20Hz à 14KHz. La résolution spatiale de la pellicule utilisée pour les pistes son optiques est d'environ 100 lignes par mm. Puisqu'une pellicule de 35 mm défile à la vitesse de 456 mm par seconde, la bande passante maximale du film ne peut dépasser 22Khz [34].

Dans une pellicule 35 mm, la piste son optique a une largeur standardisée de 2,13 mm. Dans le cas de notre défileur, la taille de la piste en largeur est de 512 pixels ; notre système offre donc une résolution horizontale de 6105 pixels par pouce.



FIGURE 1.7 – Défileur du projet RESONANCES.

Une seconde de son correspond à une longueur de 18 pouces, soit 45,72 cm. La fréquence d'acquisition de notre défileur est de 48000 lignes par seconde ; la résolution verticale est donc de 2666 pixels par pouce.

Conversion numérique en audio

Une fois la piste son numérisée, il est facile de la convertir numériquement en un signal audio. En effet, puisque l'amplitude du signal audio est, à chaque instant, proportionnelle à la quantité de lumière qui traverse la piste son, elle peut donc être mesurée par la somme des niveaux de gris de chaque ligne perpendiculaire à la direction de défilement du film. Plus cette somme est importante, plus large est la région transparente et plus la quantité de lumière traversant la piste son est importante (figure 1.8). Cette même méthode est utilisée pour les pistes à largeur variable et pour les pistes à densité variable.

Notons que la conversion numérique d'un signal audio en une piste son peut se faire en affectant à chaque valeur d'amplitude du signal, une largeur qui lui est proportionnelle (dans le cas des pistes à largeur variable) ou un niveau de gris qui lui est proportionnel (dans le cas des pistes à densité variable).

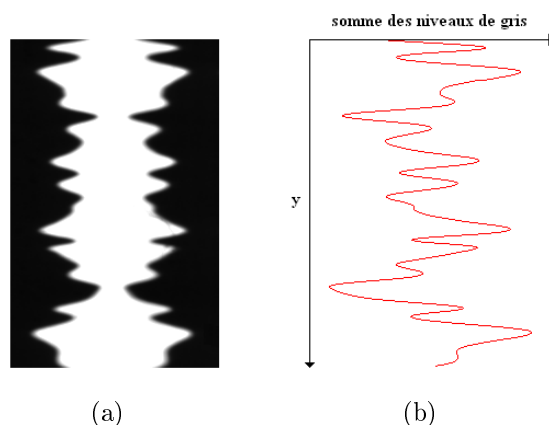


FIGURE 1.8 – (a) Piste son optique. (b) Somme des niveaux de gris de chaque ligne.

1.7 Tests d'écoute en aveugle

Si le fait de juger qu'une image soit visuellement de meilleure qualité qu'une autre, est une tâche qui n'est pas toujours évidente, comparer deux sons est une tâche qui est encore plus difficile.

Il est intéressant de noter ici que l'être-humain est capable d'entendre les fréquences allant de 20 Hz à 20 kHz¹. Certaines études ont montré que cette limite supérieure a tendance à diminuer avec l'âge ([62, 36]). D'autres études ont montré que les femmes entendent en moyenne mieux que les hommes [20].

Pour avoir une évaluation objective des méthodes de restauration, nous avons organisé une campagne de tests d'écoute en aveugle. 36 personnes (26 hommes et 10 femmes) dont l'âge varie entre 23 et 67 ans, ont participé à ces tests. L'histogramme de la figure 1.9 montre la distribution des âges de ces personnes. Tous ces participants travaillent à l'Ecole des Mines de Paris sur le site de Fontainebleau.

Ces tests d'écoute ont été effectués dans une salle fermée en utilisant une paire d'enceintes de monitoring de type *Yamaha HS50M*. Il a été demandé à chaque personne d'écouter deux fichiers audio, autant de fois qu'elle le trouvait nécessaire, puis de choisir lequel des deux lui semble être de meilleure qualité, ou de dire si aucune différence n'est perçue (fig. 1.10).

Chaque participant avait une série de 9 comparaisons à effectuer. Pour des raisons de fiabilité, chaque série de comparaison a été effectuée trois fois. L'ordre des deux fichiers audio correspondant aux pistes son avant et après restauration est attribué à chaque fois

1. Une recherche récente a montré que, bien que l'on ne soit pas capable d'entendre les fréquences au-delà de 20 kHz, leur présence engendre un changement dans l'électro-encéphalogramme des personnes les écoutant. De plus, ces mêmes personnes donnent qualitativement des jugements différents du son, selon que ces fréquences «ultrasoniques» soient présentes ou pas ([48, 47]).

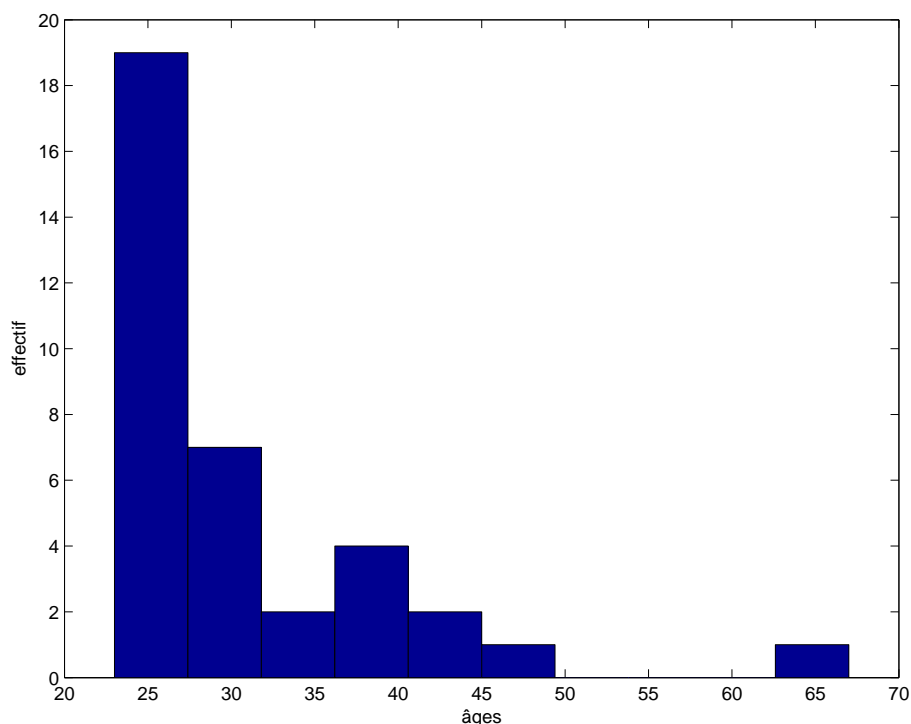


FIGURE 1.9 – Histogrammes des âges des personnes participant aux tests en aveugle.

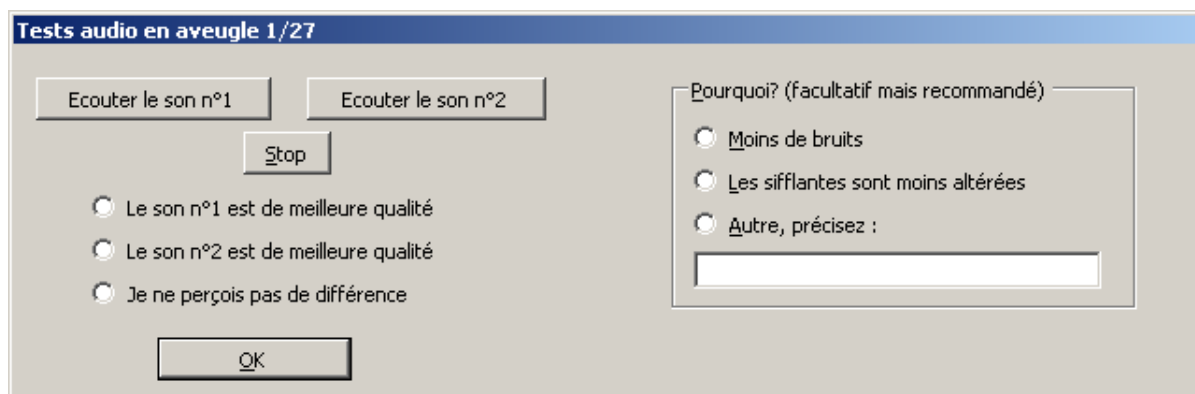


FIGURE 1.10 – Interface utilisée pour les tests audio en aveugle.

aléatoirement. La figure 1.11 montre un exemple de fichier de sortie correspondant à un participant.

Comme on peut le constater, en comparant deux sons, le jugement d'un participant n'est pas forcément le même au cours des trois comparaisons. A notre avis, ceci est dû à la fatigue de la personne après un certain nombre d'écoutes. De plus, dans la plupart des cas, ceci se produit quand la différence n'est pas perceptible, mais les participants ont

Identifiant : Philippe Sexe : Homme Age : 37 ans Durée totale du test : 45 minutes			
	Comparaisons	Nombre d'écoutes	Justification du choix
1 ^{ère} série	s1_avant > s1_apres	12,10	Sifflantes moins altérées
	s2_apres = s2_avant	8,6	-
	s3_avant < s3_apres	2,2	Sifflantes moins altérées
	s4_apres = s4_avant	13,11	-
	s5_apres = s5_avant	8,11	-
	s6_avant < s6_apres	4,3	Moins de bruit
	s7_apres = s7_avant	9,8	-
	s8_avant < s8_apres	5,6	-
	s9_apres = s9_avant	3,3	-
2 ^{ème} série	s1_apres > s1_avant	4,3	Sifflantes moins altérées
	s2_apres > s2_avant	5,4	-
	s3_avant < s3_apres	6,6	-
	s4_avant = s4_apres	5,4	-
	s5_avant = s5_apres	6,5	-
	s6_apres > s6_avant	3,2	Moins de bruit
	s7_apres = s7_avant	11,11	-
	s8_avant = s8_apres	8,7	-
	s9_avant = s9_apres	7,6	-
3 ^{ème} série	s1_avant < s1_apres	2,1	-
	s2_avant < s2_apres	6,5	Sifflantes moins altérées
	s3_apres > s3_avant	12,11	-
	s4_avant > s4_apres	6,5	Sifflantes moins altérées
	s5_apres = s5_avant	8,8	-
	s6_avant < s6_apres	2,2	Moins de bruit
	s7_avant < s7_apres	4,4	Moins de bruit
	s8_apres = s8_avant	20,20	-
	s9_apres = s9_avant	5,5	-

FIGURE 1.11 – Fichier de sortie correspondant à un participant aux tests d'écoute en aveugle. Les notations : son1>son2, son1=son2 et son1<son2, signifient respectivement que : le son1 est de meilleure qualité de le son2, qu'aucune différence n'a été perçue et que le son2 est de meilleure qualité de le son1.

tendance à trancher d'un côté ou d'un autre, parce qu'ils se disent qu'il s'agit forcément de deux sons différents. A l'issue de ces tests, seuls les jugements qui sont identiques au cours des trois séries de comparaisons seront pris en compte. A titre d'exemple, l'avis du participant dont le fichier de sortie est montré dans la figure 1.11 ne sera pris en compte qu'en ce qui concerne la 3^{ème}, la 5^{ème}, la 6^{ème} et la 9^{ème} comparaison.

Nous restons convaincus que nous ne répondons pas à toutes les recommandations concernant le bon déroulement de tels tests. Ces recommandations, publiées par «l'Union internationale des télécommunications» (UIT), concernent notamment : le choix d'auditeurs experts, la mise en place d'une procédure d'entraînement, des exigences liées aux hauts-parleurs et aux propriétés géométriques et acoustiques du local d'écoute [65]. Cependant, l'UIT affirme que ces recommandations sont à considérer dans le cas de dégradations

qui sont «si infimes qu’elles ne sont pas décelables sans contrôle rigoureux des conditions d’expérimentation» et que des essais plus simples peuvent être utilisés dans le cas de dégradations facilement décelables. Ainsi, nous pensons que la stratégie de test que nous avons suivie est largement suffisante quand les dégradations sont importantes. Cependant, dans le cas des dégradations infimes, les participants ne possèdent pas le niveau d’expertise requis pour pouvoir déceler une éventuelle différence, de plus, le local d’écoute ne répond pas aux recommandations de l’UIT. En conséquence, nous nous sommes contentés, dans de tels cas, de l’avis de nos partenaires experts du son.

1.8 Structure de cette thèse

Après ce chapitre introductif, le deuxième et le troisième chapitres abordent respectivement le problème de la restauration des pistes à largeur variable et celui de la restauration des pistes à densité variable. Le quatrième chapitre est consacré à l’étude des pistes son optiques ayant subi une mauvaise exposition. Dans le cinquième chapitre, nous décrivons la coordination entre tous les modules développés ainsi que le système global de restauration. Nous terminons ce mémoire par une conclusion générale et une liste des différentes perspectives de travaux futurs.

1.9 Conventions et notations

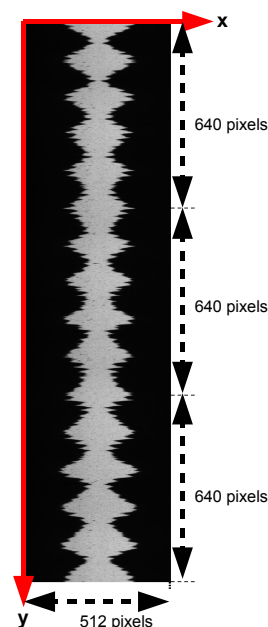
Une image f est représentée par une fonction entière d'un sous-ensemble approprié de \mathbb{Z}^2 . La dilatation et l'érosion morphologiques avec un élément structurant B seront notées respectivement δ_B et ε_B . De même, l'ouverture et la fermeture morphologiques avec B seront notées respectivement γ_B et φ_B .

On définit *l'axe des x* et *l'axe des y* qui correspondent respectivement à l'axe horizontal et à l'axe vertical (c'est-à-dire la direction du défilement du film). $f(x, y)$ représente le pixel de l'image f dont les coordonnées sont respectivement x et y .

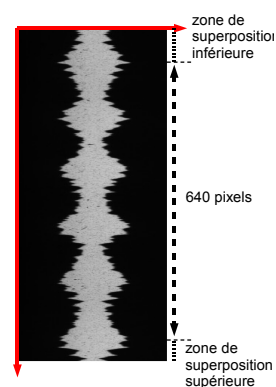
Le défileur utilisé contient une caméra linéaire permettant la capture de 48000 lignes de 512 pixels par seconde. Pour des raisons pratiques, une image est enregistrée après la capture de chaque 640 lignes. Aussi, nous entendons dans ce document par «image», un «buffer» de 512×640 pixels. Les traitements que nous allons développer s'appliqueront ainsi sur la succession de ces images (figure 1.12(a)).

Notons que pour la gestion des effets de bords, certains traitements nécessitent que chaque image soit complétée par les bords de ses deux images voisines comme le montre la figure 1.12(b). Nous parlerons alors de *zone de superposition* inférieure ou supérieure selon la position du bord ajouté par rapport à l'image traitée.

Enfin, nous entendons par «image contenant des hautes fréquences», une partie d'une piste son optique, dont le son correspondant, contient des hautes fréquences.



(a)



(b)

FIGURE 1.12 – (a) Traitement des images en série. (b) Gestion des effets de bord.

Restauration des pistes à largeur variable

Selon les estimations du Centre National de Cinématographie, environ 75% des films sonores optiques contiennent des pistes à largeur variable (appelées aussi pistes à densité fixe ou VA pour «variable area soundtracks»), contre 25% avec des pistes à densité variable (ou VD pour «variable density soundtracks»). Pour cette raison, une grande partie de notre travail a été consacrée à la restauration de ce type de pistes.

2.1 Définitions

Nous définissons dans cette section plusieurs notions liées aux pistes VA.

2.1.1 Propriétés des pistes à largeur variable

Rappelons que l'amplitude du signal audio dans les pistes VA est modulée par une région transparente généralement comprise entre deux régions opaques symétriques.

Le contour de la zone de transition (entre la région transparente et la région opaque) d'une piste VA doit être «causal». c'est-à-dire que chaque ordonnée de la piste son est associée à un seul point du contour. A titre d'illustration, le contour de la figure 2.1(a) est causal mais pas celui de la figure 2.1(b). Cette dernière ne représente donc pas une «configuration» valide pour une piste VA.

En plus de la causalité des contours, les pistes VA bilatérales doivent être symétriques par rapport à un axe de symétrie vertical. Ainsi, la figure 2.1(c) correspond à une piste son valide contrairement à la figure 2.1(d).

2.1.2 Dynamique audio

Rappelons que l'amplitude du signal audio peut être mesurée par la somme des niveaux de gris de chaque ligne. Dans le cas de notre défileur, le niveau de gris de chaque pixel

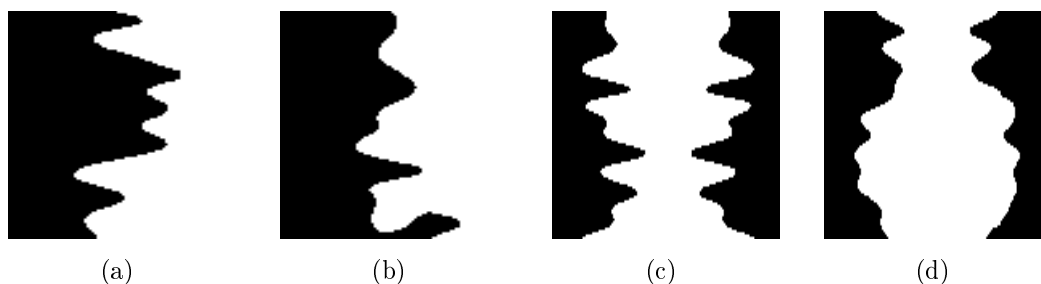


FIGURE 2.1 – (a) Contour causal. (b) Contour non causal. (c) Piste symétrique. (d) Piste non symétrique.

est codé sur 12 bits, la taille de chaque ligne étant de 512 pixels (soit 2^9), la somme des niveaux de gris de chaque ligne est donc codée sur 21 bits. Cependant, pour des raisons pratiques, les fichiers audio que l'on génère sont codés en 16 bits.

2.2 Travaux antérieurs

Très peu de travaux concernent la restauration des pistes VA. Dans cette section, nous faisons le bilan de ces travaux en montrant leurs apports et leurs limites.

2.2.1 Brevets

J. Valenzuela et V. Williams sont les inventeurs de plusieurs brevets qui portent sur le sujet ([70, 68, 67]). Ces brevets ont été déposés par «Technicolor», aujourd'hui filiale de «Thomson». Les revendications présentes dans ces brevets concernent essentiellement des dispositifs d'acquisition, d'affichage et de conversion en fichiers audio. D'un point de vue algorithmique, il est fait mention d'une restauration en utilisant un seuil ajusté manuellement [70], et d'un filtrage de la piste son en utilisant des filtres médians [67]. Dans la plupart de ces brevets, les inventions font mention de la correction d'azimut. Cette correction est cependant effectuée en alignant la caméra et non pas automatiquement.

D. Poetsch et al. sont les auteurs d'un brevet déposé par «Thomson Licensing» [50]. Il est fait mention dans ce brevet de la détection de l'axe de symétrie dans les pistes VA bilatérales.

Enfin, D. George de «Digital Technology Systems of California, Inc.» a breveté un dispositif de lecture de piste son optique permettant de s'affranchir de quelques défauts simples [17].

En conclusion, les brevets que l'on a pu énumérer proposent des dispositifs de numérisation des pistes son optiques. Les traitements proposés sont généralement rudimentaires

et ne permettent de restaurer que quelques défauts simples. Malheureusement, les traitements algorithmiques ne sont décrits que très vaguement et ne nous permettent pas d'effectuer une étude plus poussée.

2.2.2 Travaux académiques

Thèse de master de P. Streule

En 1999, P. Streule a proposé un système de restauration allant de la lecture de la piste son jusqu'à la génération du fichier audio final [59]. Le scanner qu'il utilise offre une résolution temporelle de 48500 échantillons par seconde de son. La résolution en largeur est cependant assez limitée et ne permet de générer que des fichiers audio d'une dynamique de 13 bits par échantillon.

En ce qui concerne la restauration, Streule propose un premier algorithme d'accentuation de contraste servant à réduire le grain du film. Une seconde étape permet d'éliminer les poussières en supprimant toutes les composantes connexes isolées. Streule propose aussi un algorithme pour la détection de l'axe de symétrie en détectant le milieu des zones de transition (entre les régions sombres de la piste son et la région claire). Malheureusement, les zones de transition détectées peuvent correspondre aux contours de certains défauts et non pas à ceux de la piste son, ce qui peut dérouter considérablement cet algorithme. Une dernière étape servant à corriger la symétrie de la piste est proposée : si deux pixels équidistants de l'axe de symétrie ont deux intensités différentes, cette méthode impose la symétrie en favorisant le pixel le plus clair. Cependant, comme l'affirme Streule lui-même, en cas de présence de taches claires dans les régions sombres, cette méthode génère des erreurs importantes.

Bien que Streule a le mérite d'être l'un des premiers à avoir abordé le sujet de la restauration de la piste son optique par traitement d'images, les traitements qu'il propose restent élémentaires et ne permettent de restaurer que quelques défauts simples.

Travaux de D. Richter et al.

D. Richter et al. proposent une méthode de localisation des défauts dans les pistes multiples à largeur variable [54]. Cette méthode applique un filtre passe-bas à l'image. Après binarisation, les défauts restants sont suffisamment larges pour être facilement détectés. Les auteurs ne traitent pas le problème de leur élimination.

Travaux de A. Kuiper

A. Kuiper propose une méthode de détection de taches [42]. Les taches considérées étant beaucoup plus claires que le reste de l'image, une binarisation permet donc de les isoler. Une suite d'opérations morphologiques est ensuite appliquée pour une meilleure localisation de la tache et pour la suppression des pixels isolés. Malheureusement, cette méthode suppose que la tache a un niveau de gris plus élevé que le reste de l'image ce qui n'est souvent pas le cas. Cette méthode ne peut donc pas être toujours utilisée. A. Kuiper propose aussi une méthode de correction des taches détectées dans les pistes multiples à largeur variable [43]. La piste contenant des taches est remplacée localement par la zone d'une autre piste avec laquelle elle présente le plus de similarité.

2.2.3 Méthode d'E. Brun

Durant son stage de fin d'études au Centre de Morphologie Mathématique, E. Brun a proposé une méthode de restauration des pistes son optiques à largeur variable ([15, 16]). La méthode qu'il propose commence par filtrer l'image par une ouverture par reconstruction pour réduire les défauts (fig. 2.2(b)). La quasi-distance [8] de l'image filtrée est ensuite calculée (fig. 2.2(c)). L'axe de symétrie est obtenu en calculant la régression linéaire de la ligne de crête de la quasi-distance (fig. 2.2(d)). L'image est ensuite segmentée en calculant la ligne de partage des eaux sur l'image du gradient initialisée avec trois marqueurs : l'axe de symétrie étant le marqueur central et les deux autres marqueurs correspondant aux deux lignes parallèles à l'axe de symétrie. La distance entre le marqueur central et les deux marqueurs latéraux est donnée par la valeur maximale de la quasi-distance plus une valeur empirique (fig. 2.2(e) et 2.2(f)). La dernière étape permet d'imposer la symétrie de l'image segmentée et de corriger ses contours en favorisant à chaque fois le pixel noir le plus proche de l'axe de symétrie (fig. 2.2(g)).

Le stage d'E. Brun a été effectué sur des images obtenues à partir d'un télécinéma. Ces images n'étaient pas très nombreuses et ne comportaient pas une grande variété de défauts. Les résultats obtenus avec sa méthode étaient cependant visuellement satisfaisants selon nos partenaires industriels. Après avoir testé cette méthode sur les images acquises avec la caméra linéaire, nous avons rencontré les limites suivantes :

- Mauvaise segmentation des structures fines : ce problème est dû au fait que la ligne de partage des eaux ne permet pas d'avoir une segmentation précise des structures fines.
- Limites de la quasi-distance : malgré plusieurs optimisations de cette opération [26], elle nécessite toujours un temps de calcul important. De plus, la détection de l'axe de symétrie en utilisant la quasi-distance peut donner des résultats insatisfaisants, notamment dans le cas de taches sombres situées dans les régions claires (fig. 2.3).
- Effet de la binarisation : cette opération implique une quantification importante du

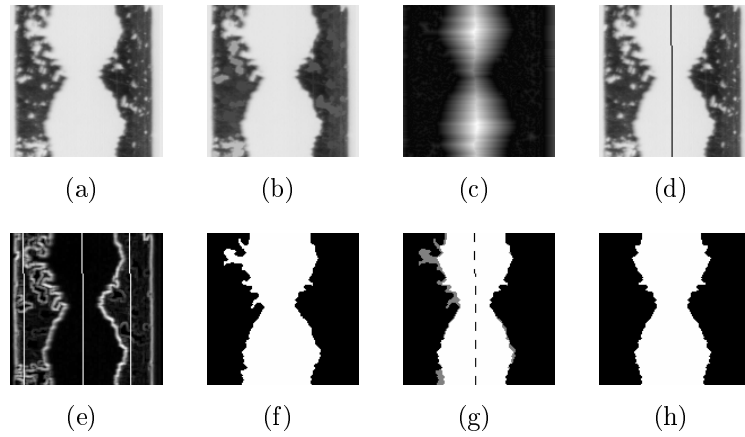


FIGURE 2.2 – Etapes de la méthode de Brun. (a) Image originale. (b) Ouverture par reconstruction. (c) Quasi-distance. (d) Axe de symétrie. (e) L'image du gradient et marqueurs. (f) Segmentation par ligne de partage des eaux. (g) Correction de symétrie et des contours. (h) Image restaurée.

signal, elle génère ainsi un bruit de fond très audible. Cet effet sera étudié plus en détail ultérieurement dans ce chapitre.

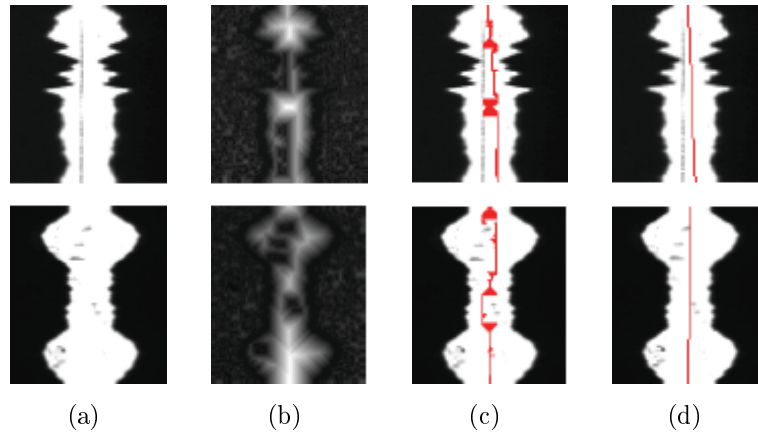


FIGURE 2.3 – Problème de la détection de l'axe de symétrie en utilisant la quasi-distance. (a) Images originales. (b) Images des quasi-distances. (c) Lignes de crête. (d) Axes de symétrie.

2.2.4 Bilan de l'étude bibliographique

Nous concluons que les méthodes existantes sur la restauration des pistes VA sont élémentaires et ne permettent de restaurer que quelques défauts simples.

Notons enfin que nous avons commencé cette thèse sur la base du travail effectué par Brun, pour cette raison, nous comparerons souvent dans la section suivante notre méthode par rapport à la sienne.

2.3 Méthode proposée

Toutes les pistes VA mises à notre disposition sont bilatérales. Nous nous sommes donc contenté d'étudier ce type de pistes. Nous verrons plus loin dans ce chapitre, dans quelle mesure nos méthodes de restauration peuvent être appliquées à la restauration des pistes unilatérales.

Notons aussi que les pistes dont nous disposons sont soit mono optiques, soit stéréo optiques. La méthode que l'on propose ici concerne uniquement la restauration des pistes mono. La restauration des pistes stéréo peut être effectuée en traitant chacune des deux pistes séparément, comme nous allons le détailler ultérieurement dans ce chapitre.

2.3.1 Détection de l'axe de symétrie

La piste son optique à largeur variable bilatérale est symétrique par rapport à un axe central. La méthode de restauration doit détecter automatiquement cet axe de symétrie pour restaurer le côté dégradé en tenant compte de l'autre côté.

T. Yuan et X. Tang ont classifié les méthodes de détection de l'axe de symétrie en deux catégories [74]. Les méthodes de la première catégorie sont basées sur les contours et les formes des objets ([75, 38], [60, 51, 73]). Ces méthodes nécessitent donc une segmentation préalable. Les contours détectés par cette segmentation pourraient correspondre aux contours de certains défauts et non pas à ceux de la piste son, ce qui déroutait considérablement ces méthodes. Les méthodes de la seconde catégorie, quant à elles, utilisent des techniques basées sur des caractéristiques de régions ([57, 24, 61, 39]). Certaines des méthodes de cette catégorie nécessitent un temps de calcul trop important pour notre tâche, les autres ne produisent pas de résultats suffisamment précis. Enfin, notre problème est différent de ceux traités dans ces articles, puisque l'on connaît approximativement la position de l'axe de symétrie mais on souhaite la détecter avec précision. Pour ces raisons, nous jugeons que les algorithmes proposés ne sont pas adaptés à notre tâche.

Nous avons décidé de calculer le point de symétrie de chaque ligne perpendiculaire à la direction de défilement du film. L'axe de symétrie sera déterminé comme étant la régression linéaire de tous ces points de symétrie. Plus précisément, soit L_y une ligne perpendiculaire à la direction du défilement du film (fig. 2.4(a)). Pour chaque pixel p de la ligne L_y , on calcule la moyenne des différences absolues entre les niveaux de gris de chaque paire de pixels dans L_y symétriques par rapport à p (fig. 2.4(b)). On appelle le

valeur obtenue le pli de la ligne L_y au pixel p .

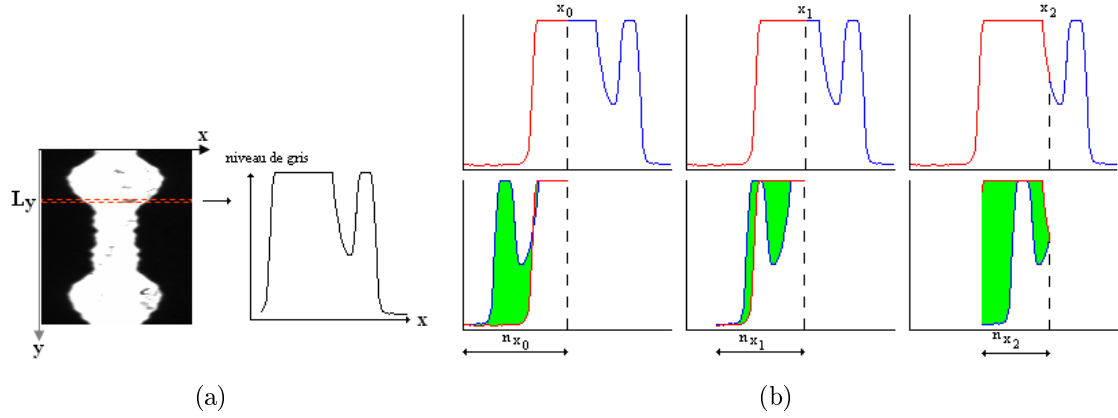


FIGURE 2.4 – Pli d’une image. (a) Profil des niveaux de gris d’une ligne. (b) Pli de cette ligne en trois pixels différents : la valeur du pli correspond à la moyenne de la région remplie.

En calculant cette valeur pour chaque pixel dans chaque ligne d’une image f , on obtient l’image pliée F :

$$F(x, y) = \frac{\sum_{k=1}^{n_x} |f(x + k, y) - f(x - k, y)|}{n_x}, \quad (2.1)$$

où n_x est le nombre de différences calculées quand on calcule le pli de la ligne L_y au pixel $p(x, y)$, c’est-à-dire le nombre de pixels dans L_y symétriques par rapport à $p(x, y)$.

Le pixel de symétrie de chaque ligne L_y est celui pour lequel les deux régions situées d’un côté et de l’autre présentent le plus de similarité ; ce pixel est donc celui qui minimise la valeur du pli, il constitue ce que l’on appellera ci-après un *point de crête* dans l’image pliée. Contrairement à la quasi-distance, l’influence des pixels sombres situés dans les régions claires sur l’image pliée est négligeable puisque le pli prend en compte toute la région et non pas les zones de transition seulement. En plus, l’algorithme du pli est moins consommateur en termes de temps de calcul que celui de la quasi-distance.

Pour déterminer l’axe de symétrie, on calcule la régression linéaire des points de crête de l’image pliée. Notons que ces points de crête sont recherchés dans un intervalle limité autour du centre de l’image. Sa largeur est empiriquement fixée à un cinquième de la largeur de l’image. En conséquence, l’image pliée est calculée seulement pour cet intervalle (fig. 2.5(a), 2.5(b)). A cause de la présence éventuelle de défauts locaux, les points de crête peuvent être déviés considérablement de l’axe de symétrie (fig. 2.5(c)). Pour que de tels points ne déroutent pas le résultat final, on considère lors du calcul de la régression linéaire, uniquement les points appartenant à une *zone de fiabilité*. Nous avons choisi de

déterminer cette zone comme étant l'intervalle de largeur 5 pixels contenant le plus de points de crête (fig. 2.5(d) et 2.5(e)). Il est à noter aussi que, si la valeur d'un point de crête (dans l'image pliée) n'est pas «suffisamment petite», ceci signifie que les deux côtés ne sont pas symétriques. Pour cette raison, lors du calcul de la régression linéaire, on considère seulement les points de crête dont la valeur (dans l'image pliée) est inférieure ou égale à un certain seuil, empiriquement fixé à 80.

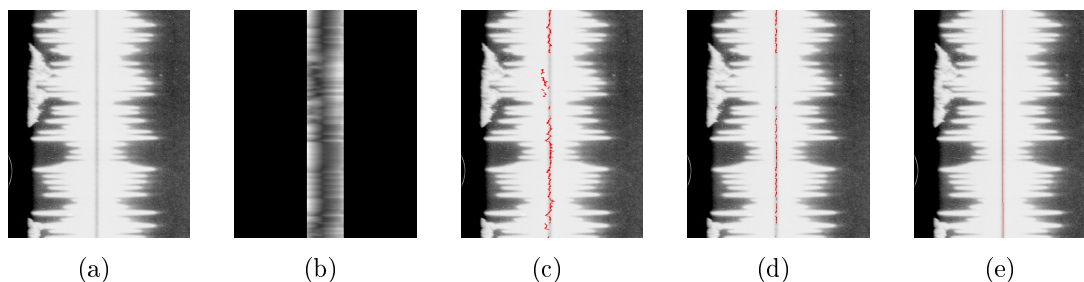


FIGURE 2.5 – Etapes de détection de l'axe de symétrie. (a) Image originale. (b) Image pliée. (c) Points de crête. (d) Points de crête fiables. (e) Régression linéaire des points de crête fiables.

Dans la plupart des cas, la position de l'axe de symétrie d'une image à la suivante, ne présente qu'une très faible variation. Pour cette raison, une fois la position de l'axe de symétrie initialisée, nous restreignons l'intervalle de calcul de l'image pliée à 5 pixels autour de la position précédente de l'axe de symétrie. Si cet intervalle ne contient pas de points de crête (ce qui ne se produit que très rarement), ceci signifie que la position de l'axe de symétrie a changé considérablement et l'algorithme sera alors réinitialisé (en calculant l'image pliée pour l'intervalle considéré initialement).

Enfin, dans le cas d'images très dégradées, la zone de fiabilité ne contient qu'un nombre très limité de points de crête. Si tel est le cas, on garde la position de l'axe de symétrie obtenue pour l'image précédente. Une image est jugée très dégradée si la zone de fiabilité contient un nombre de pixels de crête inférieur ou égal au dixième de la hauteur de l'image (fig. 2.6).

L'organigramme de la figure 2.7 récapitule les étapes de détection de l'axe de symétrie.

2.3.2 Correction de l'azimut

On parle de *déviatio*n d'azimut quand la fente de l'enregistreur optique ou la caméra linéaire n'est pas exactement perpendiculaire à la direction de défilement du film. La figure 2.8 montre une piste son optique présentant une déviation d'azimut.

Même à la fin des années 50 et au début des années 60, une déviation d'azimut de plus de $3,81 \cdot 10^{-4}$ cm pour une longueur de fente de 0,24384 cm était considérée comme

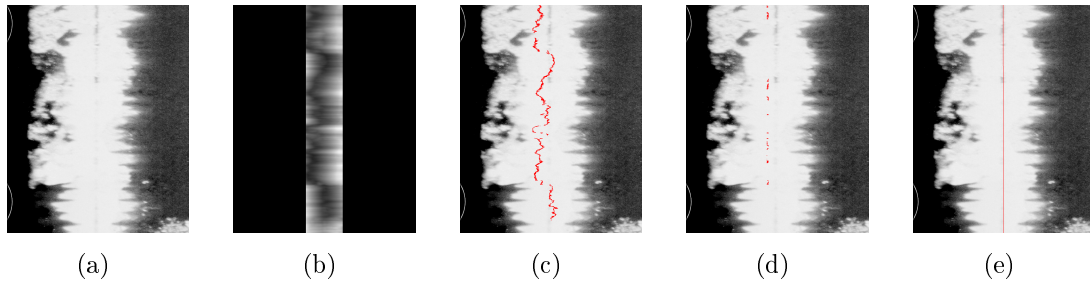


FIGURE 2.6 – Détection de l'axe de symétrie dans le cas d'une image très dégradée. (a) Image originale. (b) Image pliée. (c) Points de crête. (d) Intervalle contenant le plus de points de crête. (e) Axe de symétrie de l'image précédente conservé.

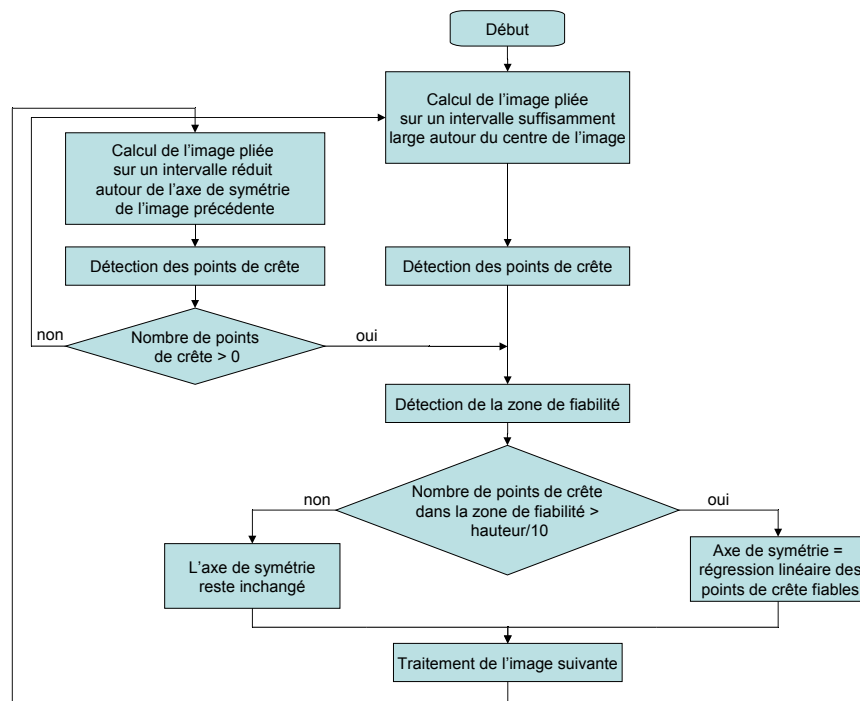


FIGURE 2.7 – Etapes de détection de l'axe de symétrie.

inacceptable ([59, 64]). Une déviation d'azimut sur les pistes VA se traduit par la réduction des hautes fréquences du signal et l'introduction de plusieurs harmoniques ([59, 27]). En considérant le cas de notre résolution, la déviation d'azimut ne doit pas dépasser 0,3 pixel pour une ligne de 512 pixels. Il est donc important de considérer les *fractions de pixel* pour avoir une bonne correction de l'azimut.

La correction d'azimut peut être effectuée par un cisaillement inverse (deskewing). Cette opération est largement utilisée pour la correction d'alignement des documents numérisés (voir à titre d'exemple [18]). Streule [59] a proposé un algorithme de correction

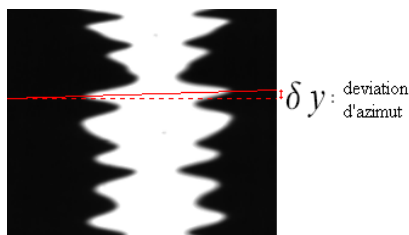


FIGURE 2.8 – Piste son présentant une déviation d'azimut.

d'azimut ; la méthode qu'il propose ne permet cependant pas de détecter automatiquement la valeur de la déviation d'azimut. De plus, l'algorithme qu'il propose n'est valable que dans certains cas particuliers comme nous allons le voir ultérieurement.

Détermination du degré de déviation

En considérant que la caméra linéaire n'est pas tout à fait perpendiculaire à la direction de défilement du film, les lignes «correctes» constituant la piste son sont des lignes en biais par rapport aux lignes capturées. Pour un angle de déviation inférieur ou égal à 45° , les pixels d'une telle ligne intersectent l'image originale en deux ou trois pixels (fig. 2.9(a)). Nous proposons de calculer les intensités de ces pixels comme étant la somme pondérée des intensités des pixels les constituant ; les coefficients de pondération correspondent aux surfaces de contribution de chaque pixel original. Streule propose un algorithme similaire pour la calcul des lignes en biais [59] ; il considère cependant que les pixels des lignes en biais sont toujours constitués de deux pixels originaux, or ceci est seulement valable pour des valeurs de déviation $\delta y = \frac{1}{N}$ où N est entier.

On considère séparément le cas de deux pixels originaux constituant le pixel de la ligne en biais (fig. 2.9(b)) et le cas de trois pixels (fig. 2.9(c)). Après quelques calculs géométriques, on obtient l'algorithme 1 permettant de calculer les intensités d'une ligne en biais pour une déviation d'une valeur δy sur une image I .

Nous cherchons maintenant à déterminer automatiquement la valeur de déviation d'azimut δy_0 . Il est évident que les lignes en biais obtenues avec la valeur de déviation que l'on cherche à déterminer sont symétriques. La position de l'axe de symétrie étant déterminée, nous calculons le pli des lignes obtenues pour chaque valeur de δy par rapport à l'axe symétrie. Comme énoncé précédemment, moins la valeur du pli est importante, plus les deux côtés de la ligne sont similaires ; la valeur de déviation correspond donc à la valeur de δy pour laquelle le pli est minimum. Pour illustrer ces propos, la figure 2.10 montre le pli de trois lignes ayant trois valeurs différentes de δy .

La courbe de la figure 2.11(b) montre la somme des plis de toutes les lignes de la piste de la figure 2.11(a) pour les valeurs de δy allant de -10 à 10 pixel avec un pas de $0,1$

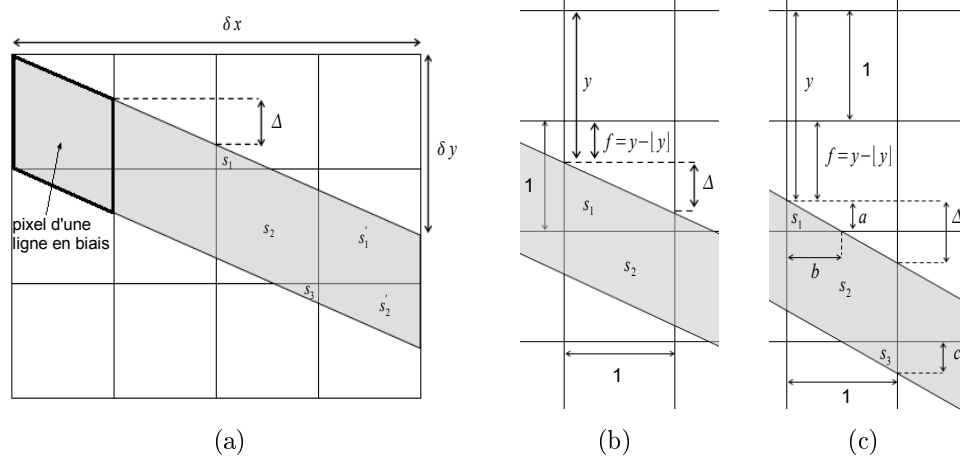


FIGURE 2.9 – (a) Ligne en biais. (b) Pixel composé des intensités de deux pixels originaux. (c) Pixel composé des intensités de trois pixels originaux.

Algorithme 1 Calcul des intensités d'une ligne en biais.

```

 $\Delta \leftarrow |\delta y| / \text{largeur}$ 
 $y \leftarrow \text{initial largeur of } y$ 
pour  $x = 1$  à  $\text{largeur}$  faire
  si  $[y] = [y + \Delta]$  alors
     $f \leftarrow y - [y]$ 
     $s_1 \leftarrow 1 - \Delta/2 - f$ 
     $s_2 \leftarrow 1 - s_1$ 
     $\text{intensité\_en\_cours} \leftarrow s_1 \cdot I(x, [y]) + s_2 \cdot I(x, [y] + 1)$ 
  sinon
     $f \leftarrow y - [y]$ 
     $a \leftarrow 1 - f$ 
     $b \leftarrow a/\Delta$ 
     $s_1 \leftarrow a \cdot b/2$ 
     $c \leftarrow \Delta \cdot (1 - b)$ 
     $s_3 \leftarrow (1 - b) \cdot c/2$ 
     $s_2 \leftarrow 1 - s_1 - s_3$ 
     $\text{intensité\_en\_cours} \leftarrow s_1 \cdot I(x, [y]) + s_2 \cdot I(x, [y] + 1) + s_3 \cdot I(x, [y] + 2)$ 
   $y \leftarrow y + \Delta$ 

```

pixel. La courbe atteint son minimum à $\delta y_0 = -2,9$ pixel.

En pratique, la valeur de déviation d'azimut est constante pour toute la piste son et ne nécessite d'être réglée qu'une fois au début du traitement. Pour avoir une mesure fiable, nous calculons la courbe en question pour une succession de 10 images.

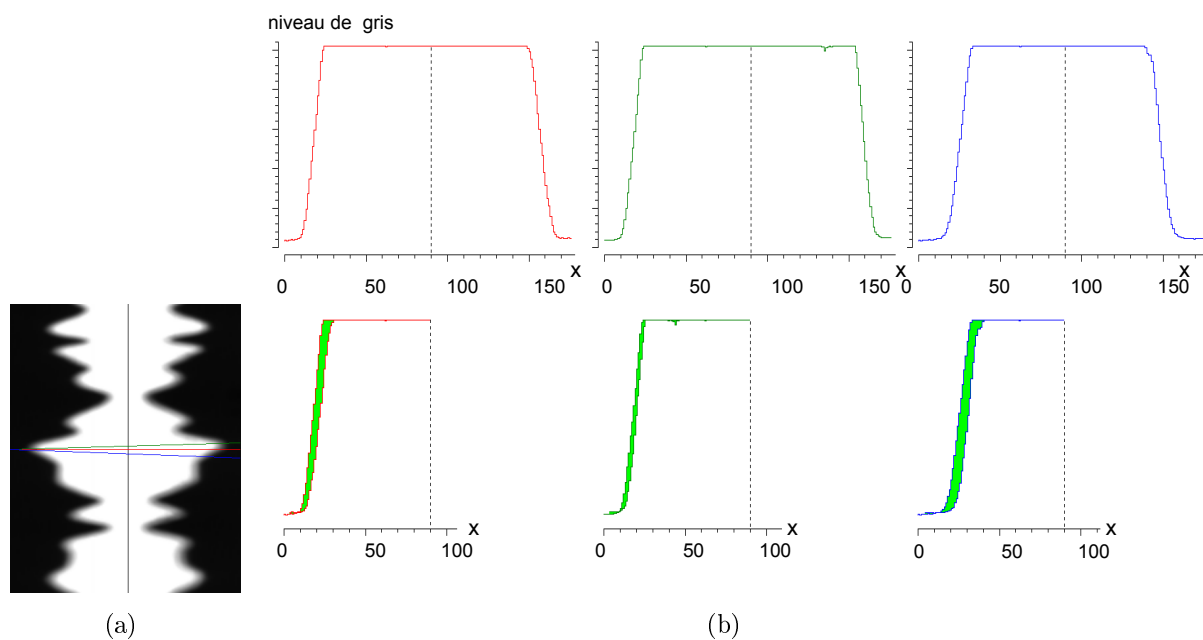


FIGURE 2.10 – Détermination du degré de déviation. (a) Trois lignes ayant des déviations différentes. (b) Pli de ces lignes par rapport à l'axe de symétrie : la valeur du pli correspond à la moyenne de la région remplie.

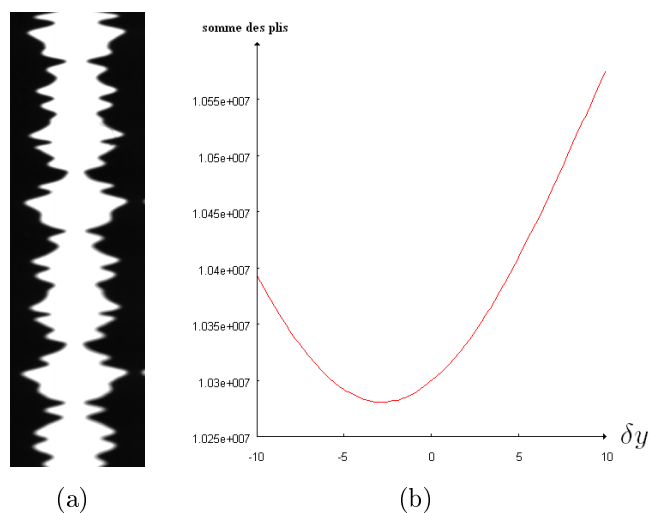


FIGURE 2.11 – (a) Image originale. (b) Somme des plis en fonction de la déviation d'azimut.

Correction de l'azimut

Nous avons initialement utilisé le même algorithme de calcul des lignes en biais pour la correction d'azimut. Nous avons ensuite remarqué que le fait d'utiliser la somme pondérée

des intensités des pixels pour le calcul des lignes en biais implique un lissage de la piste son ; ce lissage se traduit dans le domaine audio par un filtrage important des hautes fréquences. De plus, cette méthode n'est pas inversible. En effet, si on fait subir à la piste son une déviation d'azimut d'une valeur de v suivie d'une déviation de $-v$ (en utilisant cet algorithme), le résultat que l'on obtient sera une version «lissée» de l'image originale. Pour ces raisons, nous avons décidé d'utiliser, pour la correction d'azimut, l'algorithme de Bresenham [13].

Le fait d'utiliser l'algorithme de Bresenham uniquement pour la correction et non pas pour la détection est justifié par l'imprécision de cet algorithme lors de la détection de la valeur de déviation.

Notons qu'à cause des effets de bord, la correction de l'azimut nécessite une zone de superposition inférieure ou supérieure selon le signe de la déviation (voir section 1.9). La taille de cette zone est égale à $\lceil |\delta y_0| \rceil + 1$.

La figure 2.12 montre une partie d'une piste son optique avant et après la correction de l'azimut.

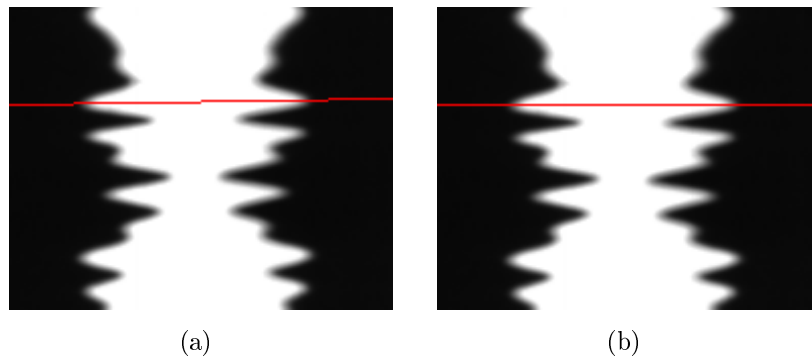


FIGURE 2.12 – (a) Image ayant une déviation d'azimut de 3 pixels. (b) Résultat de correction d'azimut. La ligne rouge a été ajoutée pour une meilleure visualisation.

Pour illustrer l'effet de cette correction dans le domaine audio, nous avons appliqué une correction d'azimut sur une piste son présentant une déviation d'azimut de 3 pixels. La figure 2.13 montre les spectrogrammes correspondant aux signaux audio avant et après la correction de l'azimut. Remarquons qu'après la correction de l'azimut, les hautes fréquences sont plus importantes, comme énoncé précédemment. L'élimination des harmoniques ne peut cependant être observée sur un tel signal à cause de la présence simultanée de plusieurs fréquences à chaque instant.

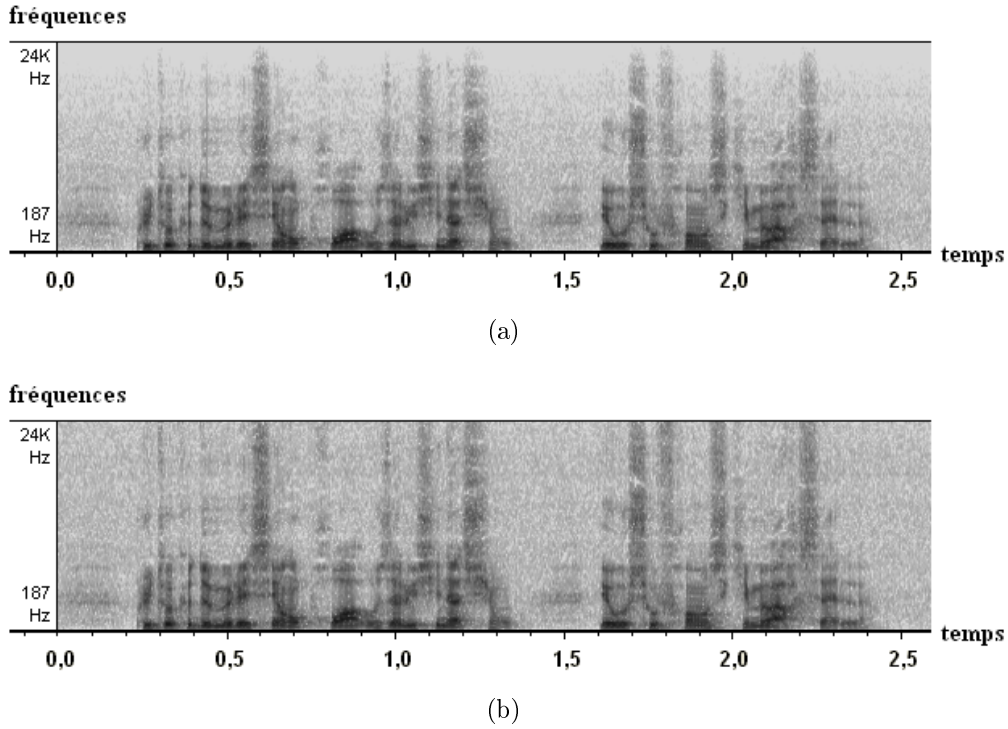


FIGURE 2.13 – (a) Spectrogramme correspondant à une piste son présentant une déviation d’azimut. (b) Spectrogramme obtenu après la correction d’azimut.

2.3.3 Renforcement de la symétrie

Le but de cette étape est de renforcer la symétrie de la piste son. Ceci est effectué en affectant simplement à chaque paire de pixels équidistants de l’axe de symétrie, la valeur minimale de leurs deux niveaux de gris, ce qui signifie que l’on favorise à chaque fois le côté le plus sombre. La figure 2.14 montre un exemple de renforcement de symétrie sur une image très dégradée.

Notons que le fait d’imposer la symétrie réduit la dynamique du signal audio de 1 bit, puisque les deux côtés de la piste sont exactement égaux et l’intégrale de chaque ligne ne peut prendre que des valeurs paires. Cette opération ne doit donc être effectuée qu’en cas de nécessité. En pratique, nous renforçons la symétrie pour une image si le volume de la différence des deux côtés situés d’un côté et de l’autre de l’axe de symétrie est supérieure ou égal à la surface de l’image multipliée par une constante, fixée empiriquement à 14. Plus formellement, la symétrie est renforcée pour une image I si :

$$S = \sum_{y=0}^{hauteur} \sum_{x=0}^{n_y} |I(a \cdot y + b - x, y) - I(a \cdot y + b + x, y)| \geq (largeur \cdot hauteur) \cdot 14, \quad (2.2)$$

où $x = a \cdot y + b$ est l’équation de l’axe de symétrie et n_{y_0} le nombre de pixels symétriques par rapport à l’axe de symétrie dans la ligne d’équation $y = y_0$. La valeur S représente la

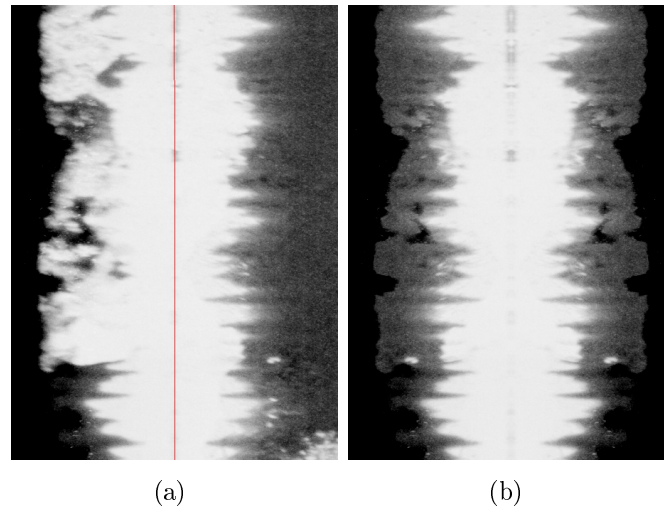


FIGURE 2.14 – Renforcement de la symétrie. (a) Image originale et axe de symétrie. (b) Résultat du renforcement de la symétrie.

somme des différences absolues des niveaux de gris de chaque paire de pixels symétriques par rapport à l'axe de symétrie. Quand sa valeur est importante, les deux côtés ne sont pas symétriques et la symétrie de l'image nécessite d'être renforcée.

Comme énoncé précédemment, cette méthode favorise à chaque fois le côté le plus sombre. Notons cependant que, le côté le plus sombre peut correspondre à la région la plus dégradée. Néanmoins, nous avons constaté que, dans la plupart des pistes son mises à notre disposition, les défauts sombres situés dans les régions claires ne génèrent pas de défauts audibles. De plus, de tels défauts ne sont pas suffisamment larges pour qu'un renforcement de symétrie soit jugé nécessaire (grâce au seuil précédemment fixé). Ainsi, cette étape ne change pas les images de la figure 2.15.

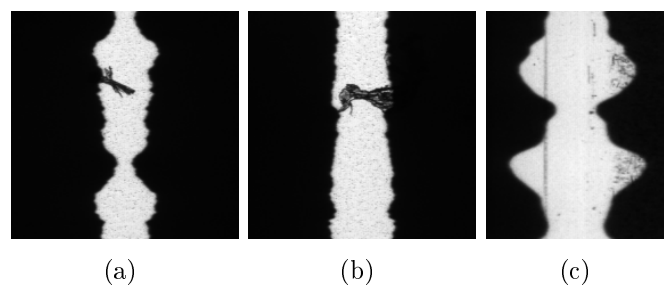


FIGURE 2.15 – Images présentant des défauts sombres dans les régions claires.

Notons enfin que le renforcement de symétrie peut être effectué après l'étape de segmentation. Son application à cette étape permet cependant d'avoir une meilleure qualité de segmentation.

2.3.4 Segmentation

L'étape la plus importante dans notre méthode de restauration est celle de la segmentation. La qualité de la segmentation est déterminante pour la qualité du résultat final. Le but de cette étape est de segmenter l'image de telle sorte que les frontières de la segmentation correspondent aux contours de la piste son. On souhaite obtenir trois régions distinctes : une région centrale claire et deux régions latérales sombres.

La segmentation ne peut pas être effectuée par un simple seuillage à cause des pixels sombres situés dans les régions claires ou vice-versa. De plus, la valeur du seuil n'est pas facile à déterminer automatiquement.

Dans la méthode de Brun [16], la segmentation était effectuée en utilisant la ligne de partage des eaux sur l'image du gradient à partir d'un ensemble approprié de marqueurs (comme décrit dans la section 2.2.3). La figure 2.16 montre une piste son, son gradient morphologique avec les différents marqueurs ainsi que la segmentation associée.

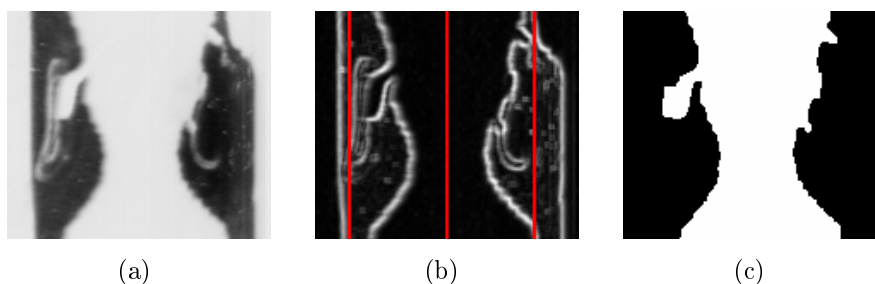


FIGURE 2.16 – Segmentation par ligne de partage des eaux. (a) Image originale. (b) Image du gradient initialisée avec trois marqueurs. (c) Résultat de la segmentation.

Il est bien connu que la ligne de partage des eaux ne permet pas une segmentation précise des structures fines. Ceci se traduit dans notre cas par une segmentation imprécise des pistes contenant des hautes fréquences. La figure 2.17 illustre ce problème.

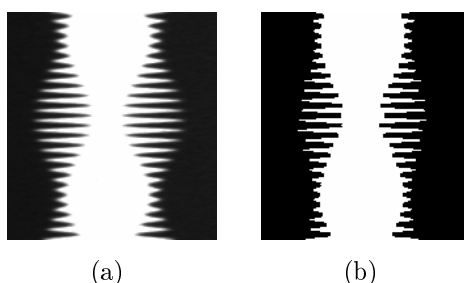


FIGURE 2.17 – (a) Image contenant des hautes fréquences. (b) Segmentation par ligne de partage des eaux.

A cause de cet inconvénient, nous avons choisi de segmenter les images en utilisant un

algorithme de croissance de régions classique [1] qui, contrairement à la ligne de partage des eaux, permet de bien préserver les structures fines de l'image. Nous initialisons l'algorithme de croissance de régions avec trois marqueurs : l'axe de symétrie étant le marqueur de la région centrale, et les marqueurs latéraux correspondant aux deux lignes situées des deux côtés de l'axe de symétrie et ayant la somme minimale des niveaux de gris.

La figure 2.18 montre quelques résultats de segmentation par croissance de régions. Remarquons que cette segmentation permet en effet une meilleure localisation des structures fines (fig. 2.18(a)). Cependant, le résultat de la segmentation n'est pas toujours causal (fig. 2.18(c)).

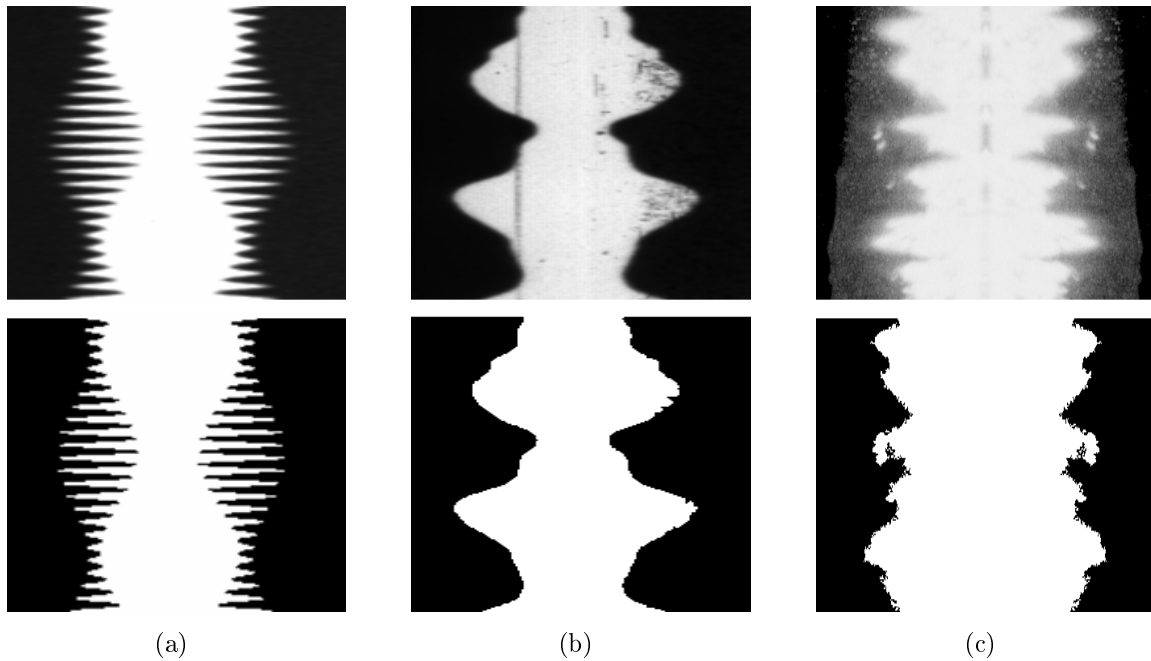


FIGURE 2.18 – Exemples de segmentation par croissance de régions.

Notons que la croissance de régions est un «opérateur global», c'est-à-dire que le résultat de segmentation d'une partie de l'image est différent selon que cette partie est traitée séparément ou dans l'image entière. Cette propriété se traduit par des discontinuités au niveau des contours en passant d'une image à la suivante. Pour éviter ce problème, il faudrait considérer des zones de superposition inférieures et supérieures suffisamment larges. Ces zones ont théoriquement une taille infinie, nous avons cependant remarqué qu'en pratique, une taille de 30 pixels ne produit pas de discontinuité. Nous avons, pour cette raison, défini pour cette étape, ainsi que toutes les étapes postérieures une zone de superposition de 30 pixels.

2.3.5 Renforcement de causalité

Le but de cette étape est d'obtenir des contours causaux. Pour ce faire, on utilise la méthode suivante : dans chaque ligne perpendiculaire à l'axe de symétrie, si un pixel noir p est situé à une distance d de l'axe de symétrie, tous les pixels sur la même ligne situés du même côté que p et à une distance supérieure ou égale à d de l'axe de symétrie deviennent noirs. La figure 2.19 montre le résultat de cette méthode sur une image segmentée.



FIGURE 2.19 – (a) Image segmentée. (b) Renforcement de causalité.

2.3.6 Lissage des contours

Bien que l'image obtenue à cette étape soit visuellement satisfaisante, le signal audio qui lui correspond n'a une dynamique audio que de 9 bits (puisque l'image est à ce stade binaire et a une largeur de 512 pixels). Ceci implique un pas de quantification important et génère donc un bruit (de quantification) plus audible que dans le signal audio original.

La figure 2.20 illustre ce phénomène. Elle montre le spectrogramme correspondant à une piste son (fig. 2.20(a)) et le spectrogramme correspondant à sa version binarisée (fig. 2.20(b)). Remarquons que dans le cas de ce dernier, les hautes fréquences sont plus importantes.

Une idée intuitive pour annuler cet effet consiste à remplacer les contours et leur voisinage dans l'image finale par une somme pondérée de l'image originale et de l'image binaire, de telle sorte que, plus le pixel est proche du contour, plus on favorise l'image originale et vice-versa.

Plus précisément, soit I_o une image d'une piste son optique (fig. 2.21(a)), I_s le résultat de sa segmentation (fig. 2.21(b)) et I_c l'image de ses contours causaux (fig. 2.21(c)). Nous calculons l'image des distances I_d à partir de I_c . Pour réduire les temps de calcul, nous utilisons une distance de chanfrein ([12, 11]) avec un masque 3×3 (fig. 2.21(d)). Pour une certaine valeur de distance d , on s'intéresse aux pixels situés à une distance du contour inférieure ou égale à d (fig. 2.21(e)). La valeur de ces pixels dans l'image finale I_f est égale à une somme pondérée de leurs valeurs dans l'image originale et dans l'image segmentée,

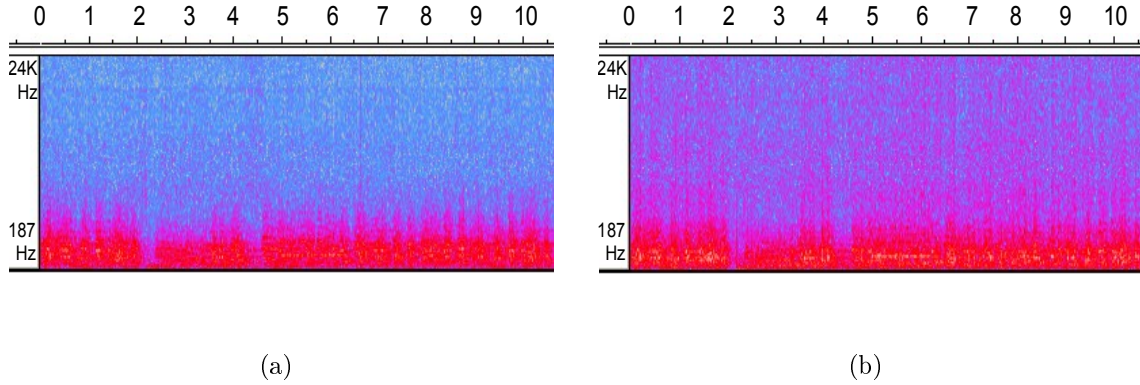


FIGURE 2.20 – (a) Spectrogramme correspondant à une piste son. (b) Spectrogramme correspondant à sa version binarisée. Le rouge représente les fréquences les plus présentes.

telle que :

$$I_f(x, y) = \left(1 - \frac{I_d(x, y)}{d}\right) \cdot I_o(x, y) + \frac{I_d(x, y)}{d} \cdot I_s(x, y). \quad (2.3)$$

Sur les pixels appartenant aux contours de l'image segmentée, la valeur de $I_d(x, y)$ est nulle ; la valeur finale d'un pixel appartenant aux contours est donc égale à sa valeur dans l'image originale I_o . A une distance d du contour, $I_d(x, y)$ est égale à d ; la valeur finale d'un pixel se trouvant à une distance d du contour est donc égale à sa valeur dans l'image segmentée I_s .

Enfin, la valeur finale d'un pixel se trouvant à une distance du contour supérieure à d est égale à sa valeur dans l'image segmentée (la formule précédente ne s'applique pas dans ce cas).

Les figures 2.21(f), 2.21(g) et 2.21(h) illustrent le résultat de cette dernière étape avec trois différentes valeurs de d .

Nous cherchons à présent à déterminer automatiquement la valeur de d que nous appelons ci-après «la taille de la zone de transition». Une «petite» valeur de d ne permet pas d'annuler complètement le bruit de fond dû à la binarisation. De même, une «grande» valeur de d introduit les éventuels défauts originellement présents dans l'image initiale. Les spectrogrammes de la figure 2.22 correspondent à des pistes son corrigées avec différentes valeurs de d .

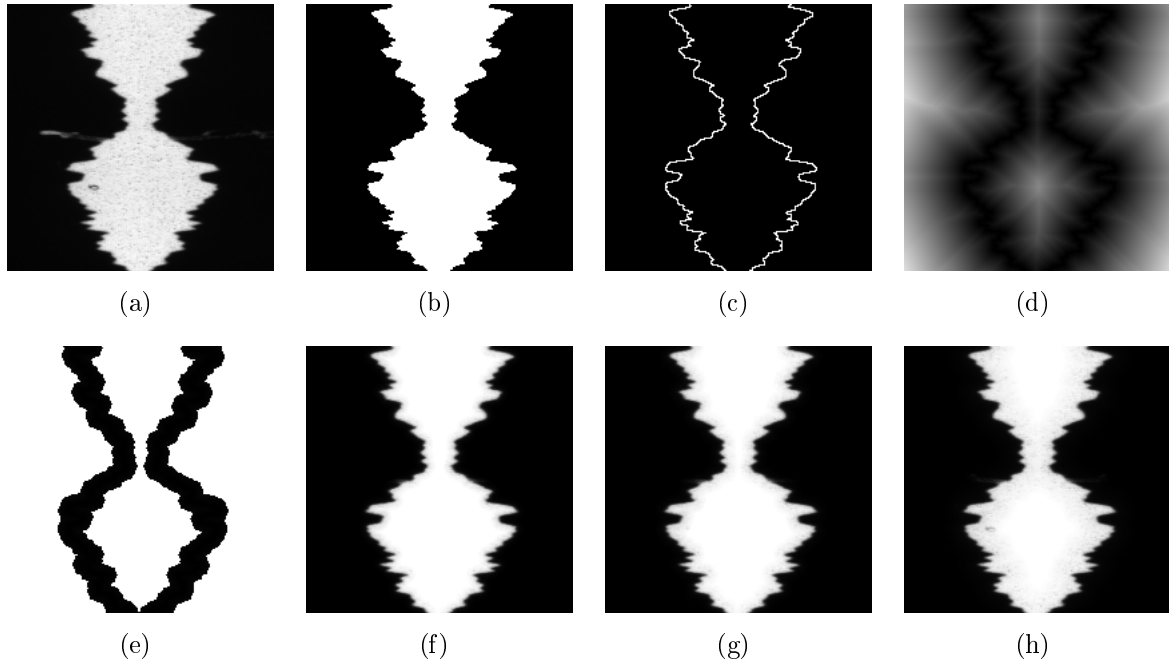


FIGURE 2.21 – (a) Image originale. (b) Image segmentée. (c) Contours de l'image segmentée. (d) Distance de chanfrein calculée sur les contours. (e) Pixels situés à une distance du contour $d \leq 7$. (f) Résultat du lissage du contour pour $d = 7$. (g) Résultat du lissage du contour pour $d = 10$. (h) Résultat du lissage du contour pour $d = 20$.

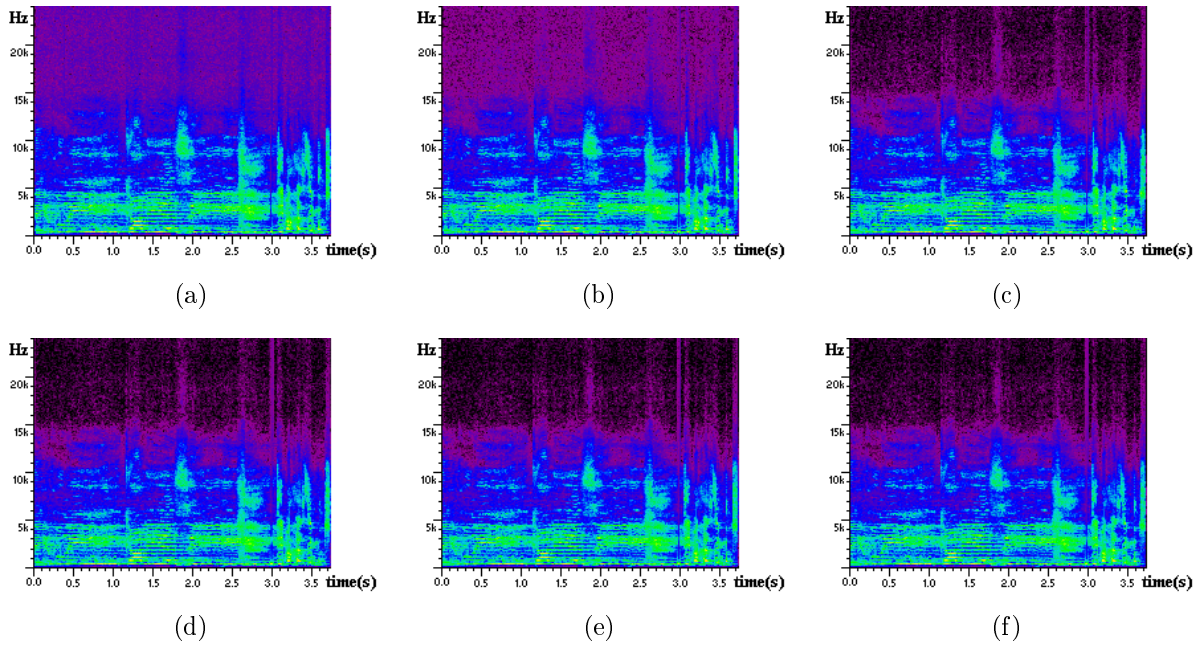


FIGURE 2.22 – Spectrogramme obtenu avec une valeur de d égale à : 0 (a), 1 (b), 2 (c), 3 (d), 4 (f), 5 (g).

Remarquons dans ces spectrogrammes que l'intensité des hautes fréquences varie en fonction de la taille choisie de la zone de transition. Pour quantifier cette variation, nous avons filtré les signaux audio générés, par un filtre passe-haut à une fréquence fixée empiriquement à 15 kHz. Nous avons ensuite calculé les énergies de chacun de ces signaux filtrés. La courbe de la figure 2.23 montre l'énergie du signal filtré en fonction de la taille de la zone de transition choisie. Cette courbe atteint son minimum à $d = 5$, au-delà de cette valeur, les hautes fréquences deviennent plus importantes à cause de l'introduction du bruit présent dans l'image initiale. Cette valeur ne change pas considérablement en fonction des pistes son ; elle varie dans l'intervalle $[5;7]$. Sa valeur peut donc être maintenue constante sans que cela affecte considérablement le résultat final.

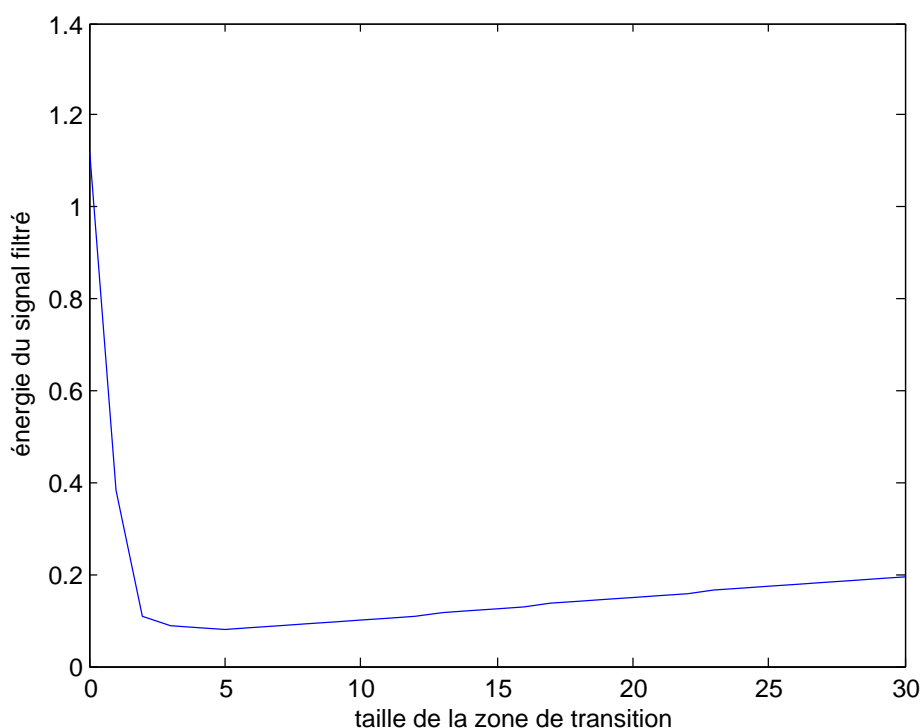


FIGURE 2.23 – Energie du signal filtré en fonction de la taille de la zone de transition.

2.3.7 Récapitulatif

L'organigramme de la figure 2.24 récapitule le schéma global de notre méthode de restauration. La communication entre les différents modules ainsi que la possibilité de leurs parallélisations feront l'objet d'un chapitre ultérieur.

La méthode développée dépend de six paramètres :

1. la longueur de l'intervalle dans lequel les points de crête sont recherchés ;

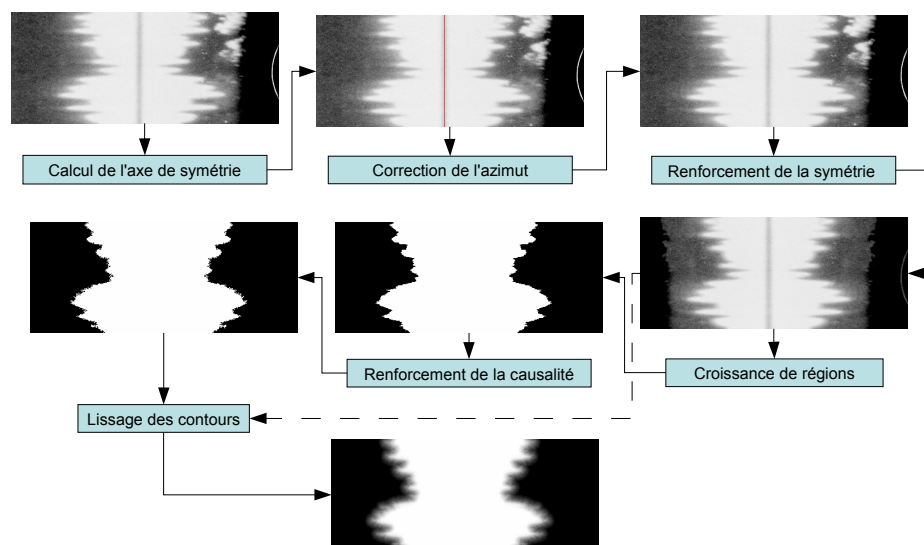


FIGURE 2.24 – Organigramme général de notre méthode de restauration.

2. le nombre de pixels de crête en dessous duquel une image est jugée comme très dégradée ;
3. la valeur minimale à partir de laquelle le point de crête est pris en considération ;
4. la différence minimale entre les deux côtés de la piste à partir de laquelle le renforcement de symétrie est jugé nécessaire ;
5. la taille des zones de superposition utilisées à partir de l'étape de segmentation ;
6. la taille de la zone de transition à prendre en compte dans l'étape du lissage des contours.

Les valeurs de ces paramètres ont été déterminées empiriquement et maintenues constantes pour toutes les pistes sur lesquelles nous avons travaillé.

2.3.8 Restauration des autres types de pistes VA

Tout le travail présenté jusqu'ici concerne la restauration des pistes à largeur variable mono optiques bilatérales. Nous montrons dans cette section, dans quelles mesures ces traitements peuvent être utilisés pour la restauration des autres types de pistes VA.

Il est difficile, selon le Centre National de Cinématographie, d'estimer avec précision la proportion des pistes son VA unilatérales. Cependant, ce genre de pistes n'a été produit qu'entre 1930 et 1935 ce qui devrait leur donner une part assez faible. Bien que nous n'ayons pas disposé dans le cadre de cette thèse de ce type de pistes, nous pensons que

les mêmes traitements peuvent y être appliqués, à l'exception de la détection de l'axe de symétrie et du renforcement de la symétrie (ce genre de pistes n'étant pas symétriques).

Les pistes VA stéréo optiques ou multiples sont, quant à elles, plus nombreuses. La restauration de ce type de pistes s'effectue d'une manière séparée sur les deux pistes composant le piste stéréo ou sur les n pistes composant la piste multiple. La position de la ligne séparatrice des pistes stéréo (ou des lignes séparatrices dans le cas des pistes multiples) est fixée par l'utilisateur.

2.4 Résultats

2.4.1 Dans le domaine image

La figure 2.25 montre quelques exemples de restauration. Remarquons que les défauts isolés sont généralement supprimés, qu'ils soient situés dans les régions claires ou dans les régions sombres. Les défauts ayant des connexions avec les contours ne sont pas complètement supprimés (fig. 2.25(b)); ces défauts ne sont cependant pas audibles, nous n'avons en conséquence pas investi davantage le problème de leur élimination. Remarquons aussi que si un seul côté de la piste son est dégradé, il est restauré par symétrie (fig. 2.25(c)). Enfin, si les deux côtés de la piste son sont sévèrement dégradés, le résultat de la restauration reste assez limité (fig. 2.25(d)). Ce dernier cas ne se présente cependant que très rarement.

Selon les archivistes du film, notre méthode de restauration est très satisfaisante, elle permet aussi bien de «nettoyer l'image que de la reconstruire par symétrie».

2.4.2 Dans le domaine audio

Comme énoncé précédemment, nous avons réalisé des tests d'écoute en aveugle pour avoir une évaluation objective de notre méthode de restauration. Les trois premières comparaisons de chaque série (voir section 1.7) correspondent à des pistes à largeur variable issues des films suivants :

1. «L'Acrobate» : un film de comédie française sorti en 1941, réalisé par Jean Boyer.
2. «Colette» : portrait de Colette, la célèbre romancière française. Court-métrage réalisé par Yannick Bellon en 1951.
3. «Délice Paloma» : réalisé par Nadir Moknèche en 2007.

La piste son de «L'Acrobate» est largement dégradée, les deux autres ne le sont que très légèrement. Le tableau 2.1 montre le résultat de ces tests.

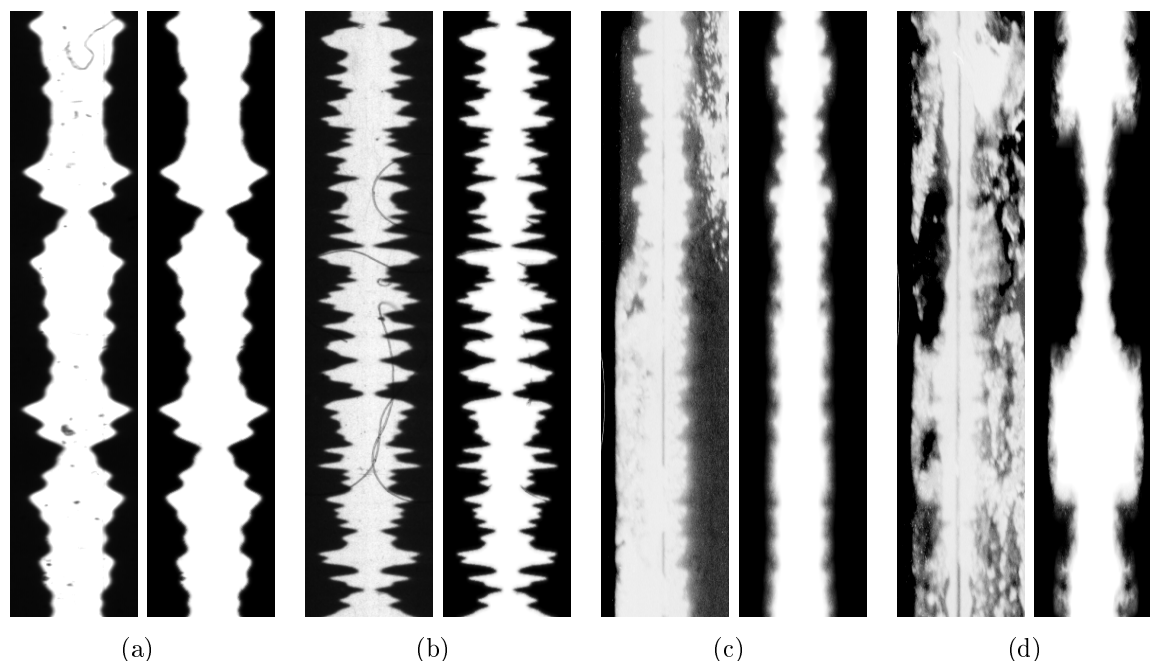


FIGURE 2.25 – Exemples de pistes son avant et après restauration : (a) Cas de défauts isolés. (b) Cas de défauts connectés aux contours. (c) Cas d'un seul côté dégradé. (d) Cas de deux côtés dégradés.

Nombre de personnes trouvant	Acrobate	Délice Paloma	Colette
le son avant restauration de meilleure qualité	2	0	0
le son après restauration de meilleure qualité	23	10	4
qu'il n'y a pas de différence	0	3	4

TABLE 2.1 – Résultats des tests en aveugle pour les pistes VA.

Nous pouvons remarquer qu'en ce qui concerne «l'Acrobate», qui correspond à la piste la plus dégradée, la plus grande majorité des participants ont préféré la version restaurée. Les 2 participants ayant préféré le son non restauré disent qu'ils avaient effectivement trouvé moins de bruit dans le son restauré, mais que le son non restauré semblait plus «naturel».

En ce qui concerne les deux autres pistes : «Délice Paloma» et «Colette», celles-ci étant moins dégradées, certaines personnes n'ont pas trouvé de différence. Cependant, une grande partie des participants a préféré la version restaurée.

Rappelons que 36 personnes ont participé à ces tests, mais plusieurs de leurs jugements ont été écartés, parce qu'ils ne sont pas identiques dans les trois séries de tests (voir section 1.7). Ainsi, dans le cas de «l'Acrobate», 11 jugements ont été écartés, 23 ont été écartés dans le cas de «Délice Paloma» et 28 dans le cas de «Colette».

Selon les archivistes du film :

- le son des cuivres est plus précis dans la version restaurée de «L’Acrobate» ;
- la version restaurée de «Colette» est une bonne démonstration de la réduction du bruit ;
- la version restaurée de «Délice Paloma» est une bonne démonstration de la réduction de distorsion de Xmodulation (ce problème sera étudié dans le chapitre «Etude des pistes son mal exposées»).

2.4.3 Temps d’exécution

Pour la restauration de 75 images (soit une seconde de son), notre méthode de restauration met environ 40 secondes, en utilisant un processeur Pentium 4 à 2,4 GHz avec 2GB de RAM. Notons que ces mêmes images étaient traitées en 300 secondes environ, avec la méthode de Brun. La différence en temps de calcul est due principalement au calcul de la quasi-distance et du gradient morphologique.

2.5 Conclusion

Nous avons présenté dans ce chapitre une méthode de restauration des pistes son optiques à largeur variable. Cette méthode comporte un algorithme robuste pour la détection de l’axe de symétrie. Des algorithmes pour la détection et la correction de la déviation d’azimut sont aussi proposés. Une segmentation par croissance de régions permet de supprimer la plupart des défauts. La méthode proposée comporte aussi d’autres étapes dont le but est de corriger la symétrie et la causalité de la piste son. Enfin, la dernière étape permet de lisser les contours pour avoir une dynamique audio égale à celle du signal original.

Les résultats obtenus sont très satisfaisants. Les images restaurées sont plus «propres» et présentent nettement moins de bruit. Dans le domaine audio, les sons restaurés sont moins bruités et de meilleure qualité. Ces avis sont partagés par les archivistes du film ainsi que la plupart des personnes ayant participé aux tests audio en aveugle.

Restauration des pistes à densité variable

Contrairement aux pistes à largeur variable, les pistes à densité variable (ou VD pour variable density soundtracks) sont peu répandues, elles ne sont plus utilisées depuis les années 50 et elles ne présentent pas plus de 25% des films sonores optiques selon les estimations du Centre National de Cinématographie. Nous parlerons dans ce chapitre du travail que nous avons effectué sur la restauration de ce type de pistes.

3.1 Définitions

Cette section regroupe quelques définitions liées aux pistes son à densité variable.

3.1.1 Propriétés des pistes à densité variable

Si dans une piste son à largeur variable, chaque partie de la piste est soit totalement opaque soit totalement transparente, dans une piste son à densité variable, chaque ligne perpendiculaire à la direction de défilement du film a la même densité, mais cette densité change le long de la piste son. L'apparence de la piste VD sera un ensemble de lignes, perpendiculaires à la direction de défilement du film. Chacune de ces lignes a un niveau de gris constant. Grâce à cette propriété, les défauts sont facilement localisables sur ce genre de pistes (fig. 3.1).

Une fois la piste son numérisée, nous pouvons observer que le niveau de gris d'une même ligne n'est pas tout à fait constant, même si cette piste est en très bon état. Ceci est essentiellement dû, d'une part, aux différences entre les sensibilités des capteurs constituant la caméra linéaire, et d'autre part, au grain du film.

3.1.2 Dynamique audio

Le procédé de conversion numérique en audio des pistes VD étant identique à celui des pistes VA, le son correspondant à ces deux types de pistes a la même dynamique audio.

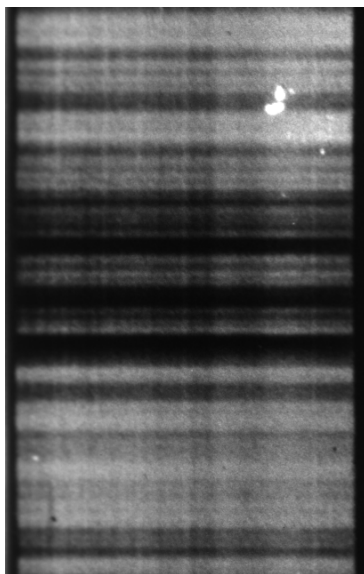


FIGURE 3.1 – Piste VD présentant un défaut.

Notons cependant que si par une quelconque méthode de restauration, on remplace tous les niveaux de gris d'une même ligne par une valeur constante, la profondeur du fichier audio sera réduite à 12 bits (qui est égale à la dynamique des niveaux de gris de l'image).

3.2 Travaux antérieurs

Les travaux concernant la restauration des pistes VD par traitement d'images sont encore moins nombreux que ceux qui concernent la restauration des pistes VA. Nous énumérons ci-dessous la liste de ces travaux.

J. Valenzuela a breveté plusieurs opérations liées à la numérisation et aux traitements des pistes son à densité variable [66]. La correction d'azimut est revendiquée dans ce brevet, mais elle est effectuée en changeant l'alignement de la caméra et non pas numériquement. Le brevet revendique aussi le fait d'éliminer les pixels dont le niveau de gris est supérieur à un certain seuil, mais aucune méthode n'est proposée pour déterminer automatiquement la valeur du seuil en question.

Richter et al. proposent une méthode de restauration des pistes VD [53]. Cette méthode commence par corriger les défauts dûs aux imperfections des capteurs (ou «fixed pattern noise» en Anglais). Ce type de corrections est actuellement implémenté de manière câblée dans les caméras et n'est donc plus nécessaire [4]. Les auteurs proposent aussi une méthode statistique pour la restauration de chaque ligne numérisée : en analysant l'histogramme des niveaux de gris de chaque ligne perpendiculaire à la direction de défilement du film,

la valeur du niveau de gris le plus récurrent est déterminée. Un seuil inférieur et supérieur sont fixés autour de cette valeur. Les pixels ayant un niveau de gris compris entre ces deux seuils restent inchangés, les autres pixels sont remplacés par le niveau de gris moyen de tous les pixels compris entre les deux seuils. Malheureusement, dans le cas de défauts très contrastés (tels que celui de la figure 3.2), le niveau de gris le plus récurrent ne correspond pas au «niveau de gris correct» de la ligne en question mais au niveau de gris du défaut. De plus, nous avons testé cette méthode sur nos bases de données, et avons remarqué qu'elle génère un bruit de fond dû aux changements importants d'intensité d'une ligne à la suivante.

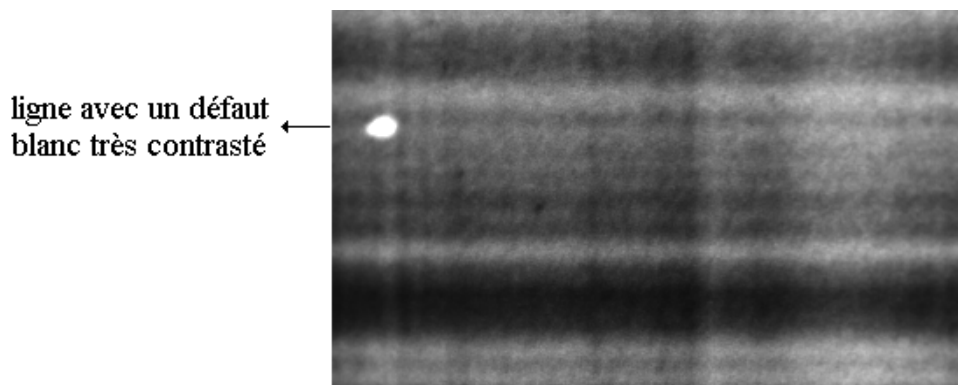


FIGURE 3.2 – Piste VD ayant un défaut très contrasté.

Notons enfin que Streule a aussi abordé le problème de la restauration des pistes VD [59]. Cependant, le seul aspect qu'il traite est celui de la correction d'azimut. Comme énoncé précédemment, cet algorithme n'est valable que dans certains cas particuliers et ne peut pas être toujours utilisé.

Nous concluons que les méthodes existantes ne proposent que des solutions simples ne pouvant faire face qu'à quelques défauts simples. Pour le traitement des défauts plus complexes que l'on a pu observer sur nos bases de données, une nouvelle méthode doit être développée.

Nous décrirons dans la section 3.3 les bases de données sur lesquelles nous avons travaillé et dans la section 3.4 la méthode que l'on a développée pour la restauration de ce type de pistes.

3.3 Bases de données

Deux pistes son VD ont été utilisées dans cette étude. Elles proviennent, toutes les deux, d'anciens films français. Nos partenaires du *Centre National de Cinématographie*

n'ont malheureusement pas pu tracer l'origine de ces pistes. La première contient un passage musical (fig. 3.3(a)) et la seconde une discussion (fig. 3.3(b)). Malheureusement, les défauts ne sont pas présents tout le long de ces deux pistes son, ils correspondent généralement à des infimes fractions de seconde, et sont rares et très espacés. De ce fait, leur répercussion dans le domaine sonore est quasi-négligeable. Pour cette raison, nous avons généré numériquement une piste correspondant à une conversation dans un restaurant, sur laquelle nous avons ajouté des défauts simulés en plaçant des formes aléatoires à des endroits aléatoires (fig. 3.3(c)). Contrairement aux pistes numérisées, les défauts des pistes simulées sont très fréquents : de l'ordre de 250 défauts ajoutés par seconde de son. En conséquence, en comparant le signal audio correspondant à la piste son simulée, avant et après l'ajout des défauts, on remarque un important bruit de fond.

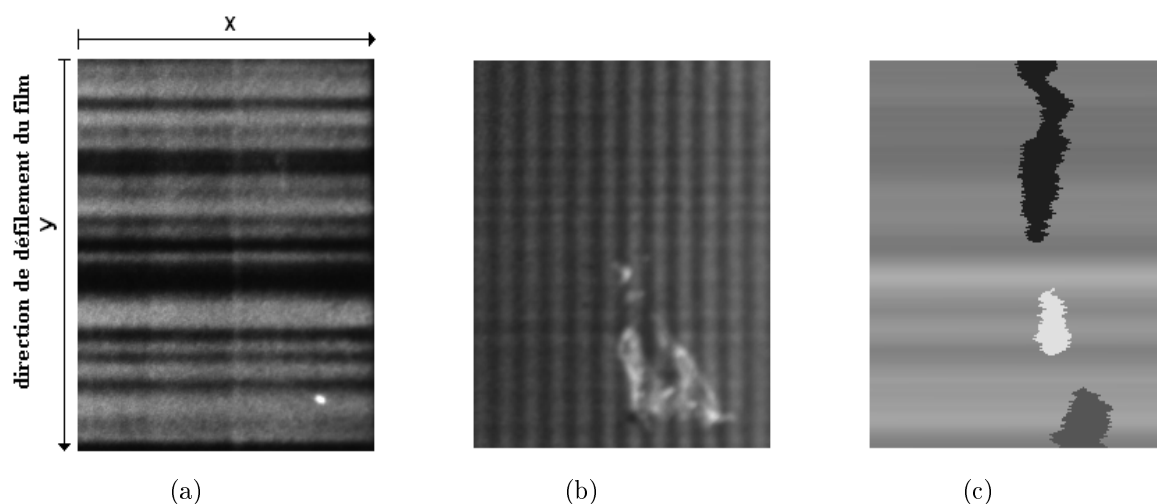


FIGURE 3.3 – Coupes des pistes VD utilisées dans cette étude.

Remarquons que la seconde piste son est striée verticalement. La raison en demeure inconnue. Les experts du «GTC» supposent cependant que ces stries sont dues à une grille placée pendant la copie pour contrôler l'exposition.

3.4 Méthode proposée

La méthode que nous proposons pour la restauration des pistes VD comporte deux étapes :

- correction d'azimut,
- correction des autres défauts.

3.4.1 Correction d'azimut

La méthode de calcul des lignes en biais est similaire à celle qui concerne les pistes VA (voir section 2.3.2).

Puisque tous les pixels d'une ligne perpendiculaire à la direction de défilement du film sont supposés avoir le même niveau de gris dans une piste VD, on cherche la valeur de déviation δy pour laquelle les intensités des pixels ont des niveaux de gris aussi proches que possible. Cette valeur de δy est celle qui minimise la variance de ces intensités.

Malheureusement, à cause de la présence éventuelle de défauts locaux, cette mesure peut donner des résultats erronés. Elle nécessite donc d'être moyennée sur une section relativement «longue» de la piste son. En pratique, la valeur de déviation d'azimut est déterminée sur une succession de 10 images.

Nous avons donc calculé pour chaque valeur de δy allant de -5 à 5 pixel avec un pas de $0,1$ pixel, la somme des variances des intensités obtenues pour 6400 lignes. La figure 3.4 montre les résultats obtenus pour les premières 6400 lignes de la seconde piste son. La valeur $\delta y = -2,8$ pixel minimise la somme des variances et sera utilisée pour la correction de l'azimut de cette piste. En procédant d'une manière similaire pour la première piste son, le résultat obtenu est $\delta y = -0,5$ pixel.

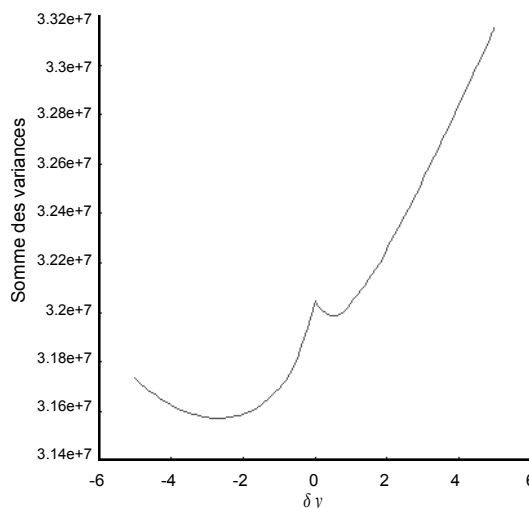


FIGURE 3.4 – Somme des variances en fonction de la valeur de déviation.

Une fois la valeur de déviation déterminée, la correction est effectuée en utilisant l'algorithme de Bresenham. En apportant cette correction, nous avons remarqué que les hautes fréquences deviennent effectivement plus significatives. Pour illustrer ces propos, la figure 3.5 montre les spectrogrammes correspondant à une partie de la deuxième piste son, avant et après la correction d'azimut.

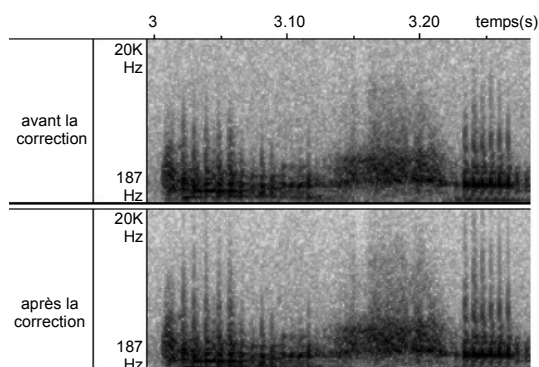


FIGURE 3.5 – Spectrogrammes avant et après la correction d’azimut. Remarquons qu’après la correction, les niveaux de gris correspondant aux hautes fréquences sont plus sombres (les hautes fréquences sont donc plus significatives).

3.4.2 Correction des autres défauts

Les résultats obtenus peuvent contenir quelques défauts locaux. La figure 3.6 montre quelques exemples des différents défauts rencontrés.

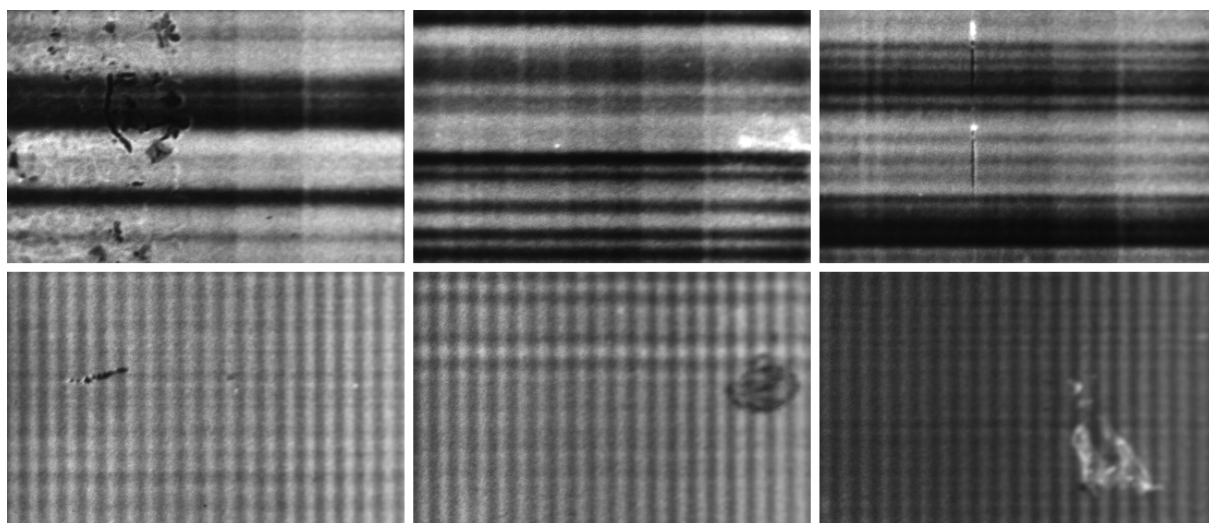


FIGURE 3.6 – Exemples de défauts sur une piste VD.

Puisque les niveaux de gris des défauts sont généralement différents du niveau de gris moyen de chaque ligne, nous avons décidé de procéder comme suit :

Pour chaque ligne perpendiculaire à la direction de défilement du film, nous calculons la moyenne des niveaux de gris v . Soit t un seuil tel que, tous les pixels ayant un niveau de gris supérieur à $v + t$ seront remplacés par $v + t$ et tous les pixels ayant un niveau de gris inférieur à $v - t$ seront remplacés par $v - t$. Les niveaux de gris résultants seront donc

compris dans $[v - t, v + t]$ (figure 3.7).

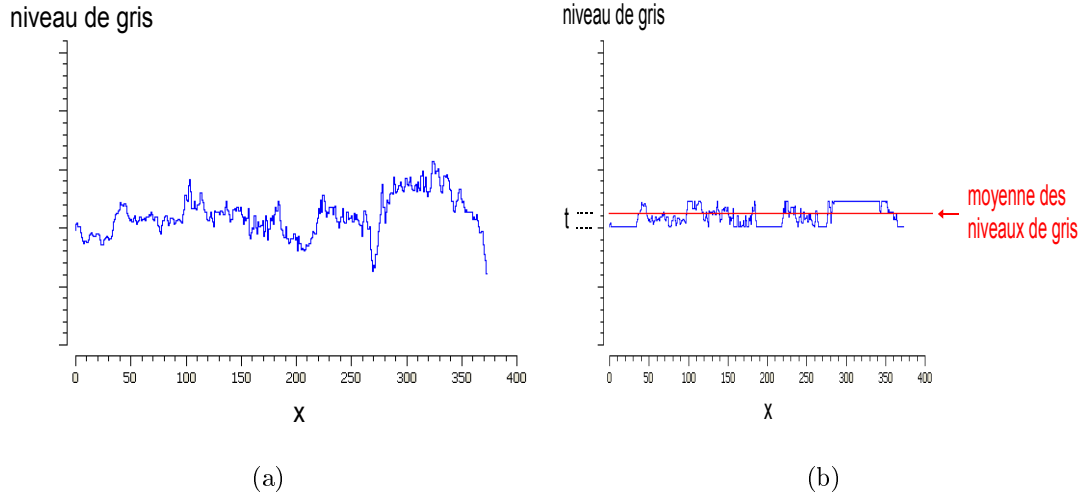


FIGURE 3.7 – Profile d’une ligne avant et après correction.

Il nous reste maintenant à déterminer la valeur de t automatiquement. Notons que pour chaque valeur de t , le signal audio résultant sera différent. Si on fixe t à une «petite» valeur, la dynamique des niveaux de gris de chaque ligne sera très réduite, le signal audio résultant sera donc très bruyé. Si au contraire, on le fixe à une «grande» valeur, seulement les pixels ayant un niveau de gris très différent de la valeur moyenne seront modifiés et le résultat final ne présentera pas une amélioration significative. Nous avons choisi de garder la valeur de t pour laquelle le signal audio résultant contient le moins de bruit et donc qui minimise la variation totale [49].

D’une manière plus formelle, pour déterminer automatiquement la valeur de t nous procédons, pour chaque valeur de t , comme suit :

- Soit $L_{y,t}$ une ligne «seuillée» à t (en utilisant la méthode précédemment décrite).
- Nous calculons pour chaque ligne $L_{y,t}$ la somme des niveaux de gris $somme_{y,t}$.
- Nous calculons ensuite la variation totale des sommes obtenues :

$$variation_totale = \sum_{y=1}^N |somme_{y,t} - somme_{y-1,t}|$$
(N étant le nombre de lignes, empiriquement fixé à 6400 lignes).
- La valeur de t que l’on retient est celle qui minimise la variation totale.

La courbe de la figure 3.8 montre la relation entre la variation totale et la valeur du seuil pour les 6400 premières lignes de la première piste son. Le seuil qui minimise la variation totale est 15. La même courbe montre un comportement similaire pour la seconde piste son et pour la piste son simulée. Elle atteint cependant son minimum à 58 sur la seconde piste son à cause de la dynamique relativement large des niveaux de gris dans cette piste, et atteint son minimum à 13 sur la piste son simulée.

Nous avons remarqué que la valeur de t qui minimise la variation totale ne varie pas le long d’une même piste son, il suffit donc de la déterminer une fois au début du traitement.

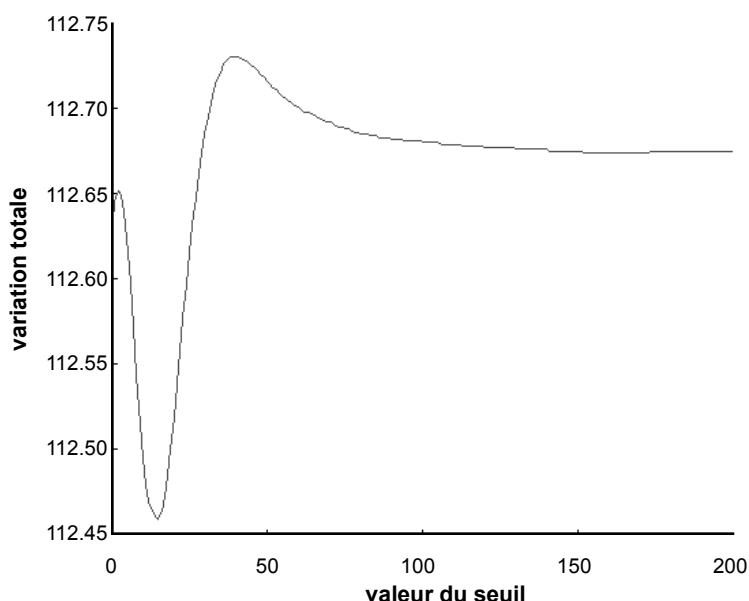


FIGURE 3.8 – Variation totale en fonction de la valeur du seuil.

Remarque

Notons qu'une dernière étape permettant de remplacer tous les pixels d'une même ligne par leur moyenne sur 16 bits¹ pourrait être considérée pour produire des images visuellement plus satisfaisantes (fig. 3.9). Cette étape n'a cependant pas d'influence sur le signal audio résultant. Nous avons, en conséquence, jugé inutile de l'inclure dans notre méthode de restauration.

3.5 Résultats

Cette méthode utilise des traitements efficaces permettant une restauration en temps réels de la piste son. Nous discutons dans cette section, des résultats de cette restauration, aussi bien dans le domaine image, que dans le domaine audio.

3.5.1 Dans le domaine image

La figure 3.10 illustre des exemples de restauration pris des trois pistes son sur lesquelles nous avons travaillé. Remarquons que les défauts sont généralement supprimés ou

1. Rappelons que si l'on remplace tous les pixels d'une même ligne par leur moyenne sur 8 bits, ceci impliquera une réduction de la dynamique du signal audio et introduira un bruit de fond très audible.

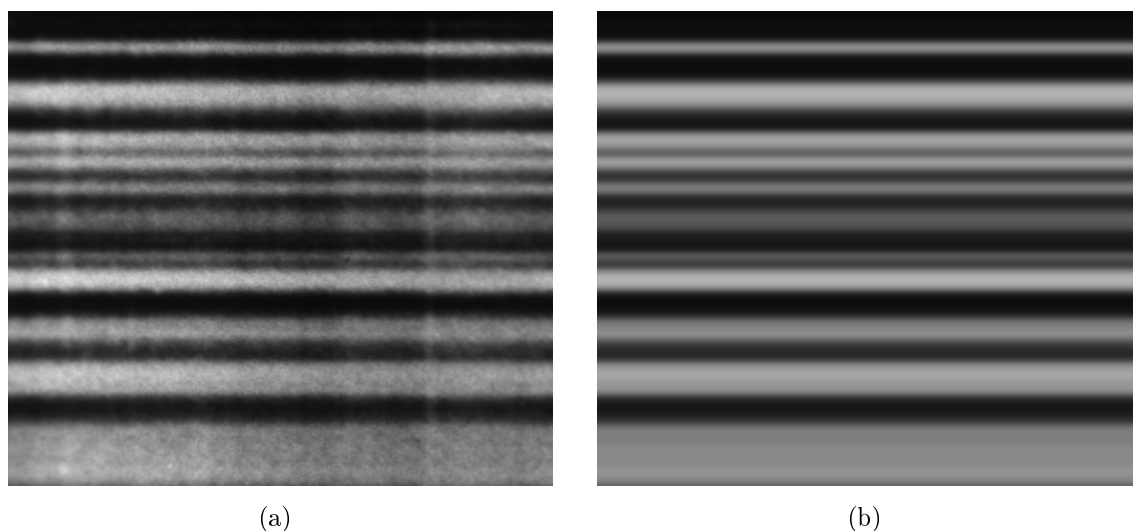


FIGURE 3.9 – Dernière étape : remplacement de tous les pixels d’une même ligne par leur moyenne sur 16 bits.

au moins réduits. Les stries de la seconde piste son ne sont pas éliminées. Nous pensons que, sans connaissance à priori sur lesquelles des stries sombres ou claires doivent être privilégiées, cette tâche ne peut être effectuée.

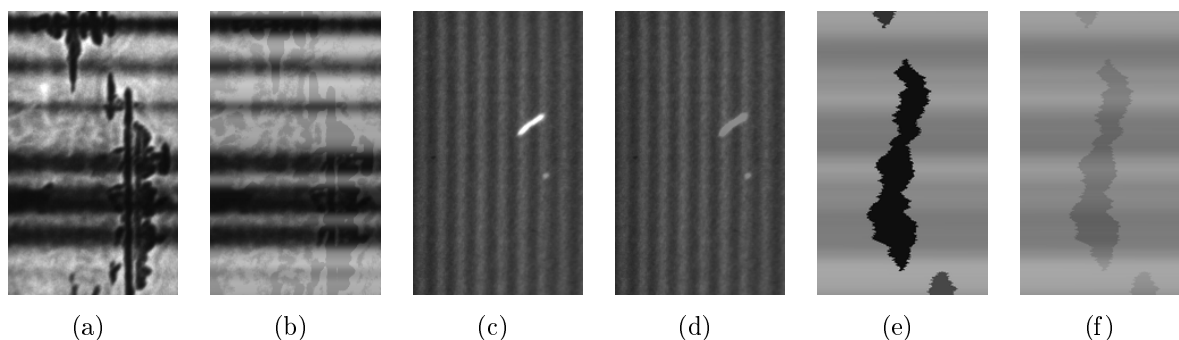


FIGURE 3.10 – Exemple pris de la première piste son avant restauration (a) et après restauration (b) . Exemple pris de la seconde piste son avant restauration (c) et après restauration (d) . Exemple pris de la piste son simulée avant restauration (e) et après restauration (f) .

Cette méthode ne permet pas de restaurer les images dans lesquelles la surface du défauts est plus importante que le reste de l’image (figure 3.11). Ce type de défauts ne se présente cependant que très rarement. Nous avons, en conséquence, jugé inutile de nous consacrer à son étude plus profondément.

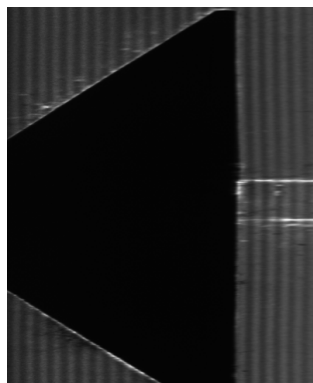


FIGURE 3.11 – Défaut étendu sur une grande surface, dû à l’assemblage de deux pistes.

3.5.2 Dans le domaine audio

Dans les deux pistes son numérisées, les défauts ne sont pas présents tout au long de la piste son, mais localisés seulement sur des régions de tailles assez limitées ; leurs effets dans le domaine audio est donc négligeable. Nous avons pu tout de même constater que la méthode de restauration préserve la qualité du signal audio et n’introduit pas de fréquences parasites (contrairement à la méthode de Richter). Cependant, sur la piste simulée, nous avons remarqué une réduction importante du bruit dû aux défauts ajoutés. Le spectre fréquentiel de cette piste son avant et après restauration montre en effet une réduction des hautes fréquences correspondant au bruit (fig. 3.12).

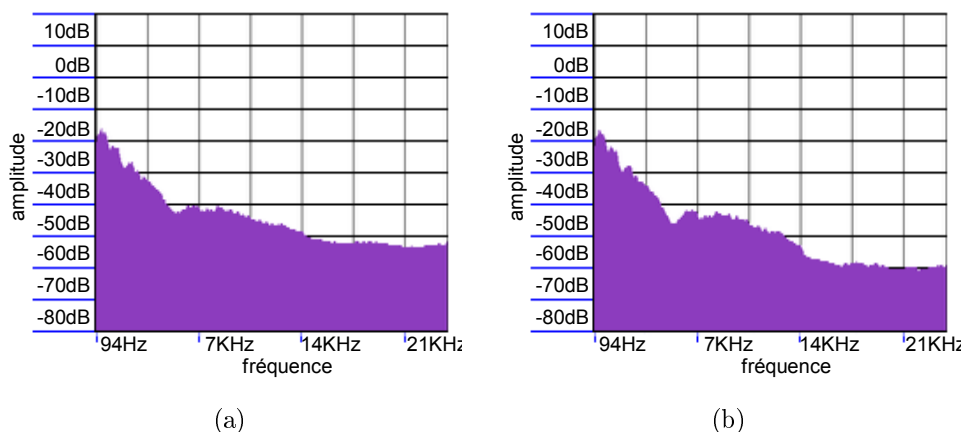


FIGURE 3.12 – Spectre fréquentiel de la piste son simulée avant restauration (a) , après restauration (b) .

Pour avoir une évaluation plus objective, nous avons réalisé sur ces mêmes pistes, des tests audio en aveugle. Elles correspondent à la 4^{ème}, 5^{ème} et 6^{ème} comparaisons de chaque série de tests (voir section 1.7).

Le tableau 3.1 montre le résultat de ces tests.

Nombre de personnes trouvant	1 ^{ère} piste	2 ^{ème} piste	Piste simulée
le son avant restauration de meilleure qualité	0	1	0
le son après restauration de meilleure qualité	0	0	34
qu'il n'y a pas de différence	19	14	0

TABLE 3.1 – Résultats des tests en aveugle pour les pistes VD.

Remarquons qu'en ce qui concerne les deux pistes numérisées (1^{ère} et 2^{ème}), presque toutes les personnes n'ont pas trouvé de différence entre la version restaurée et la version dégradée. Ceci est dû à la raison évoquée précédemment : les défauts ne sont pas présents tout le long des pistes son ; ils correspondent à des fractions de seconde difficiles à percevoir. Cependant, en ce qui concerne la piste simulée, qui elle, est beaucoup plus dégradée, toutes les personnes ont préféré la version restaurée.

Les archivistes du films ont, eux aussi, préféré la version restaurée, en ce qui concerne la piste simulée, mais n'ont pas trouvé de différences au niveau des autres pistes pour la même raison évoquée.

3.6 Conclusion et perspectives

Nous avons proposé une nouvelle méthode de restauration des pistes à densité variable. Cette méthode offre un algorithme précis de détection d'azimut et une méthode de seuillage basée sur la minimisation de la variation totale du signal audio résultant, permettant ainsi une paramétrisation automatique de la méthode.

Au niveau image, les défauts sont généralement supprimés ou réduits par notre méthode, sauf ceux qui s'étendent sur de larges zones. Au niveau audio, l'efficacité de notre méthode a été validée sur une piste son simulée. Les sons restaurés sont, en effet, moins bruités, comme en témoignent les tests d'écoute en aveugle et les avis des archivistes du film. Cependant, les pistes numérisées sur lesquelles nous avons travaillé n'étant pas très dégradées, cette méthode devra être testée sur d'autres pistes présentant des dégradations plus importantes.

Enfin, notre méthode traite les défauts dans chaque ligne séparément. La prise en compte de la distribution spatiale des défauts pour les traiter d'une manière globale dans l'image, mérite aussi d'être étudiée.

Etude des pistes son mal exposées

On s'intéresse ici à l'étude des pistes son ayant subi une mauvaise exposition. Ce phénomène est scientifiquement très riche, il a des effets aussi bien dans le domaine image que dans le domaine audio. Nous parlerons dans ce chapitre des différentes méthodes que nous avons développées pour la détection et la restauration des pistes son ayant subi cet effet.

4.1 Définitions

Dans cette section, nous définissons plusieurs notions relatives à la mauvaise exposition des pistes son.

4.1.1 Origines et effets de la mauvaise exposition

Similairement aux images du film, la piste son optique fait l'objet de plusieurs copies, allant de la piste son originelle (produite par l'enregistreur cinématographique) jusqu'à la copie finale. En conséquence, il est important de régler correctement l'exposition pour chaque copie.

Pour bien comprendre le problème de la mauvaise exposition, rappelons que dans une pellicule cinématographique, on peut observer plusieurs niveaux d'opacité (ou de transparence) connus sous le nom de «densité optique». Lors de l'enregistrement du film, c'est la valeur de la lumen¹ reçue qui se traduit par une densité optique plus ou moins importante. Contrairement à ce que l'on peut penser, la variation de cette densité en fonction de la lumen ne se fait pas d'une manière linéaire, mais selon une fonction appelée «fonction de transfert du film» (Density vs Exposure). Son allure est montrée dans la figure 4.1. Remarquons que la partie centrale de cette courbe est quasi-linéaire. Pour obtenir une exposition correcte, il faudrait alors s'assurer que c'est bien cette partie de la fonction de transfert du film qui est utilisée ([22, 44]).

1. La lumen (en anglais exposure) est une grandeur physique mesurant la quantité d'éclairement reçu par une surface pendant un certain temps, son unité de mesure est le lux seconde (lux.s), son symbole est H.

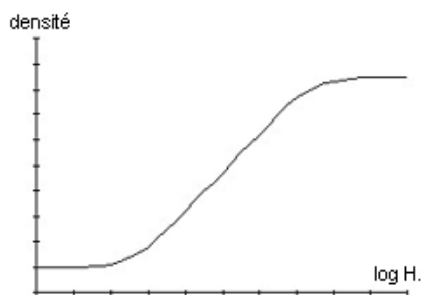


FIGURE 4.1 – Fonction de transfert du film.

Avec nos partenaires industriels, nous avons jugé inutile d'étudier le problème de la mauvaise exposition sur les pistes à densité variable, pour les raisons suivantes :

- Les pistes à densité variable présentant un problème de mauvaise exposition sont très rares, voire inexistantes dans les Archives Françaises du Film.
- Les enregistreurs optiques permettant de générer de telles pistes ne sont plus disponibles. Il n'est, en conséquence, pas possible de les reproduire.

A cause de la diffusion lumineuse, les pistes à largeur variable peuvent être altérées significativement. Dans le domaine image, l'effet d'une exposition incorrecte sur ce genre de pistes, est que les pics deviennent plus pointus et que les vallées deviennent plus creuses ou vice-versa (figure 4.2). Notons que, puisque la copie s'effectue généralement sur les négatifs, la sur-exposition rend l'image plus sombre, et non pas plus claire, contrairement à ce que son nom peut laisser entendre (et inversement pour la sous-exposition).

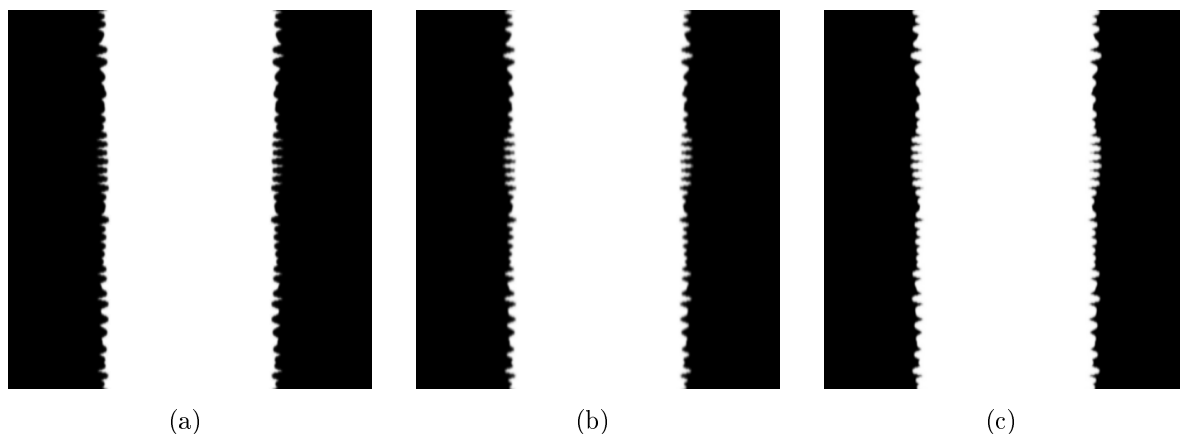


FIGURE 4.2 – (a) Cas de sur-exposition. (b) Cas d'exposition normale. (c) Cas de sous-exposition.

Dans le domaine audio, l'effet de la mauvaise exposition est négligeable sur les basses fréquences. Il est essentiellement perceptible dans les sifflantes (s, sh, ch, z...etc) et les explosives (b, p et t) [40].

4.1.2 Test de cross-modulation

La «cross-modulation» est une piste de test conçue spécialement pour vérifier l'exposition (figure 4.3). Elle comprend un signal à haute fréquence (env. 6 KHz) modulé par un signal à basse fréquence (env. 400 Hz).

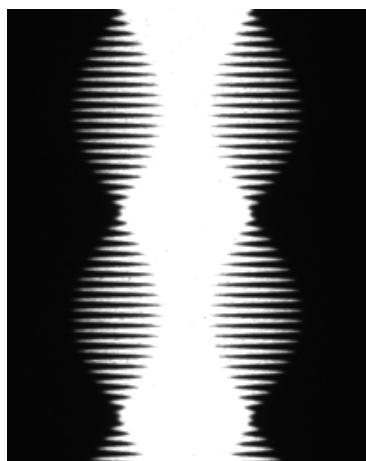


FIGURE 4.3 – Exemple de piste de cross-modulation.

La figure 4.4(a) (respectivement figure 4.4(b)) montre le signal audio correspondant à une piste de cross-modulation correctement exposée (respectivement sur-exposée). En appliquant à ces deux signaux des filtres électriques (passe-bas de 400Hz), le signal résultant est quasi-nul dans le cas correctement exposé mais pas dans le cas mal exposé (figure 4.4(c)). C'est en procédant ainsi que le technicien règle l'exposition de la piste son [40].

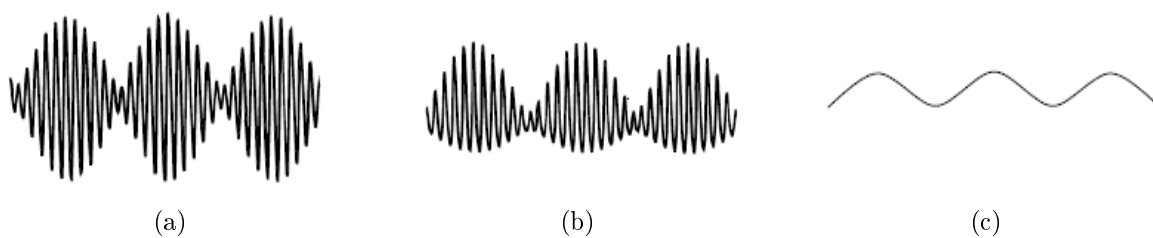


FIGURE 4.4 – (a) Signal de cross-modulation correctement exposé. (b) Signal de cross-modulation sur-exposé. (c) Filtrage du signal de cross-modulation sur-exposé.

Notons qu'un test similaire appelé «test d'intermodulation» est utilisé pour régler l'exposition dans les pistes à densité variable [27].

4.1.3 Simulation de la mauvaise exposition

Le phénomène physique causant une mauvaise exposition est bien connu et peut être simulé avec précision. Pour ce faire, on applique à la piste son optique correctement exposée (figure 4.5(a)), un filtre de convolution gaussien (figure 4.5(b)) modélisant la fonction d'étalement du point [37]. Le résultat ainsi obtenu est ensuite saturé en utilisant une fonction de type sigmoïde² (figure 4.5(c)) qui simule la fonction de transfert du film [9]. Selon la partie de la courbe utilisée, on obtient soit une sur-exposition (figure 4.5(d)) ou une sous-exposition (figure 4.5(e)).

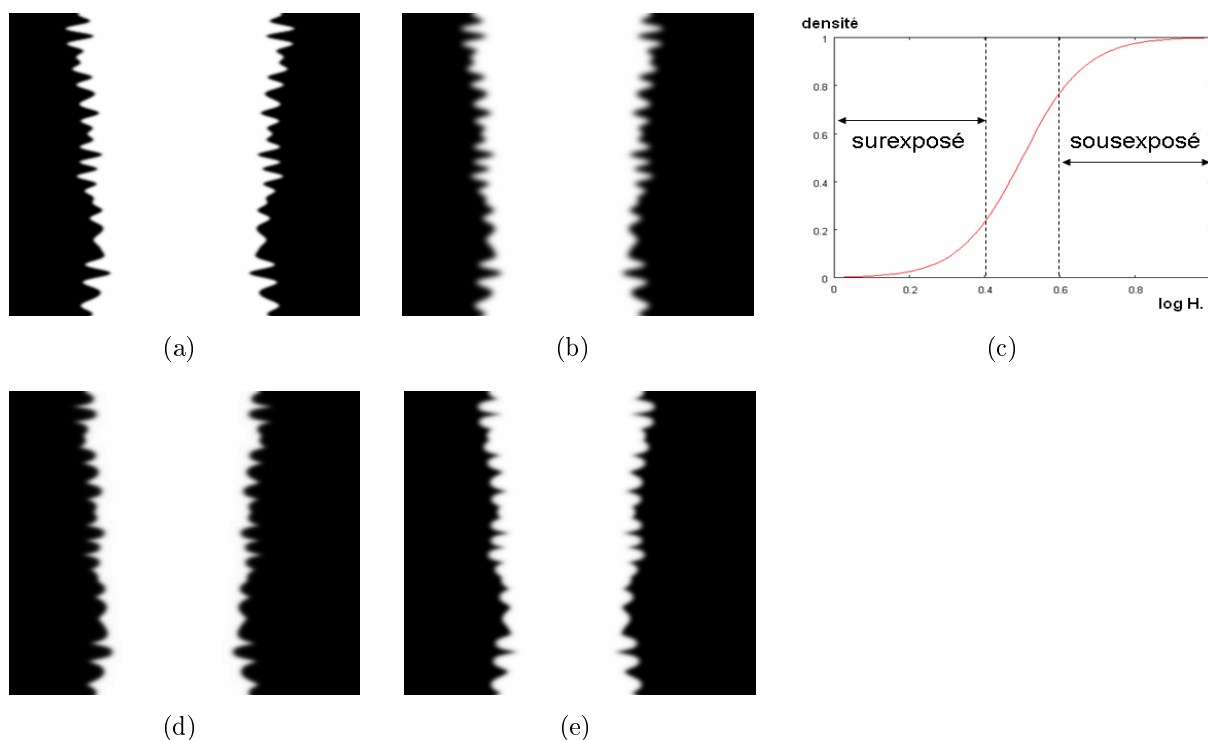


FIGURE 4.5 – (a) Image originale. (b) Gaussienne. (c) Courbe sigmoïde. (d) Image sur-exposée. (e) Image sous-exposée.

4.2 Travaux antérieurs

J. Valenzuela fait mention dans le brevet [69] de la restauration de la mauvaise exposition en utilisant des érosions morphologiques. Cependant, le procédé n'est décrit que très vaguement et ne nous a pas permis d'étudier la méthode proposée plus en détail.

Nous n'avons pas trouvé de publications portant directement sur la détection ou sur la

2. La forme générale d'une fonction sigmoïde est $P(t) = \frac{1}{1+e^{-t}}$.

correction des pistes son optiques mal exposées. Cependant, certains travaux concernent des techniques de correction d'exposition d'images en général et peuvent servir dans le cas particulier de notre étude (voir [9, 45]). Les méthodes proposées dans ces travaux décomposent l'image en un certain nombre de blocs. L'exposition est ensuite corrigée pour chaque bloc en appliquant sur les niveaux de gris de l'image une table de correspondance (LUT) basée sur la fonction de transfert du film. La correction est apportée de manière à mettre le plus en évidence les détails de chaque bloc. Nous verrons, plus loin dans ce chapitre, dans quelle mesure, ce genre de correction d'exposition pourrait être appliqué aux pistes son optiques.

Enfin, notons que le problème de la mauvaise exposition est aussi de nature géométrique, et que la morphologie mathématique [55] offre une palette d'outils bien adaptés pour traiter ce genre de problèmes. Pour cette raison, plusieurs traitements que l'on propose dans ce travail appartiennent à cette discipline.

4.3 Constitution de la base de données expérimentale

Pour étudier la mauvaise exposition, nos partenaires ont produit des pistes son correspondant à différentes expositions de la bande annonce de «Délice Paloma» et d'une piste son de cross-modulation. Les différents niveaux d'exposition ont été obtenus en faisant varier, soit la vitesse du défilement du film pendant la copie des pistes son, soit la puissance de la lampe de tirage. Le tableau 4.1 montre la liste de tous les niveaux d'exposition produits et l'identifiant affecté à chaque niveau d'exposition.

Vitesse de défilement (pieds par minute)	Puissance de la lampe de tirage (W)	Type de piste	Identifiant
120	18	sur-exposée	surex120_18
240	13	sur-exposée	surex240_13
240	18	sur-exposée	surex240_18
480	13	normalement exposée	normalex480_13
960	13	sous-exposée	sousex960_13
960	8	sous-exposée	sousex960_8

TABLE 4.1 – Différentes expositions générées

Notons que les images sont indexées de telle sorte que, deux images ayant le même indice (et deux identifiants différents) correspondent à une même partie de la piste son (avec deux expositions différentes). Les trois images de la figure 4.2 par exemple, ont le même indice (mais trois identifiants différents).

Nous avons décidé d'utiliser lors des comparaisons, les trois séquences *surex240_13*, *normalex480_13* et *sousex960_13* parce qu'elles correspondent aux trois types d'exposition (normales, sur-exposées et sous-exposées) et elles ont toutes été produites avec la

même puissance d'éclairage.

Pour avoir des niveaux de mauvaises expositions plus importants, nous avons généré numériquement la piste son correspondant au son «S» sur laquelle nous avons simulé une sur-exposition et une sous-exposition en utilisant la méthode décrite dans la section 4.1.3.

Toutes ces bases d'images qui ont été, soit capturées, soit générées numériquement, serviront dans les différentes méthodes de détection et de restauration qui font l'objet des deux sections suivantes.

4.4 Détection de la mauvaise exposition

Le degré de mauvaise exposition peut atteindre plusieurs niveaux. Au sein d'une même piste son optique, il est très probable que ce niveau soit relativement constant, puisque toute la piste est issue de la même série de copies. Cependant, l'effet d'une mauvaise exposition n'est pas le même sur toute la piste : il est important sur les hautes fréquences et négligeable sur les basses. Il est alors important de détecter le degré de mauvaise exposition pour y appliquer ensuite une correction adéquate.

Notons que la détection du degré de mauvaise exposition peut se faire dans le domaine audio [63]. Dans ce travail, nous avons développé plusieurs méthodes pour détecter la mauvaise exposition directement dans le domaine image.

4.4.1 Invariance par rapport à l'ouverture (ou la fermeture) morphologique

On a vu précédemment que l'effet d'une sous-exposition est que les pics deviennent plus pointus et les vallées plus creuses. Ceci laisse à penser que ce phénomène est similaire à une dilatation morphologique (voir figure 4.6).

Pour vérifier cette hypothèse, nous avons étudié l'invariance des pistes sous-exposées par rapport à l'ouverture morphologique. En effet, une image ayant subi une dilatation par un certain élément structurant est invariante par rapport à l'ouverture par le même élément structurant et tout élément structurant d'une taille inférieure. A cause du bruit présent dans les images, nous n'avons pas pu mettre en évidence cette invariance sans passer par une segmentation préalable.

Les images sur lesquelles nous avons travaillé dans cette étude ne présentent pas d'importantes dégradations ; un simple seuillage est donc généralement suffisant pour les segmenter. Le traitement des pistes son présentant à la fois des problèmes de mauvaise exposition ainsi que d'autres types de dégradation sera abordé dans le chapitre suivant.

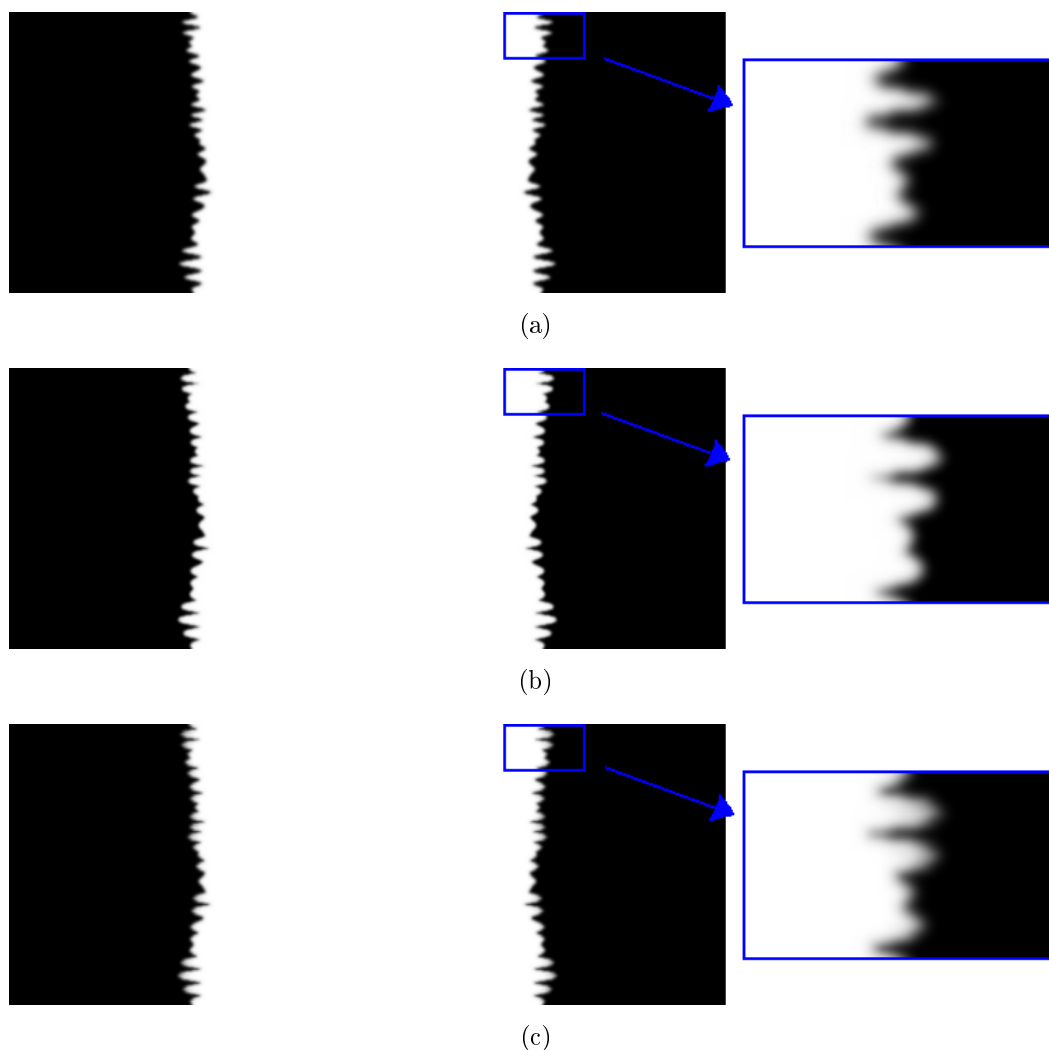


FIGURE 4.6 – (a) Image originale. (b) Image sous-exposée. (c) Image dilatée par un élément structurant de type croix.

Une valeur de seuil de 100 donne généralement une segmentation satisfaisante. Nous revenons plus loin dans ce chapitre sur l'influence de la valeur du seuil sur les différentes méthodes de détection développées. Les images ainsi segmentées pouvant contenir des «petites» régions sombres dans les zones claires ou vice-versa, nous supprimons, pour cette raison, toutes les parties connexes ayant un nombre de pixels inférieur à une certaine valeur fixée empiriquement à 200 pixels.

Nous calculons sur ces images segmentées, la succession d'ouvertures avec des éléments structurants losanges³ de tailles croissantes. Nous calculons pour chaque taille d'élément

3. Un élément structurant de type disque serait plus approprié, mais le losange fournit une bonne approximation des disques de petites tailles pour un coût de calcul moins important

structurant, le volume normalisé⁴ de l'image de différence entre l'image binaire et son ouverture. La courbe de la figure 4.7(b) illustre ce volume en fonction de la taille des éléments structurants utilisés pour l'ouverture de l'image 4.7(a).

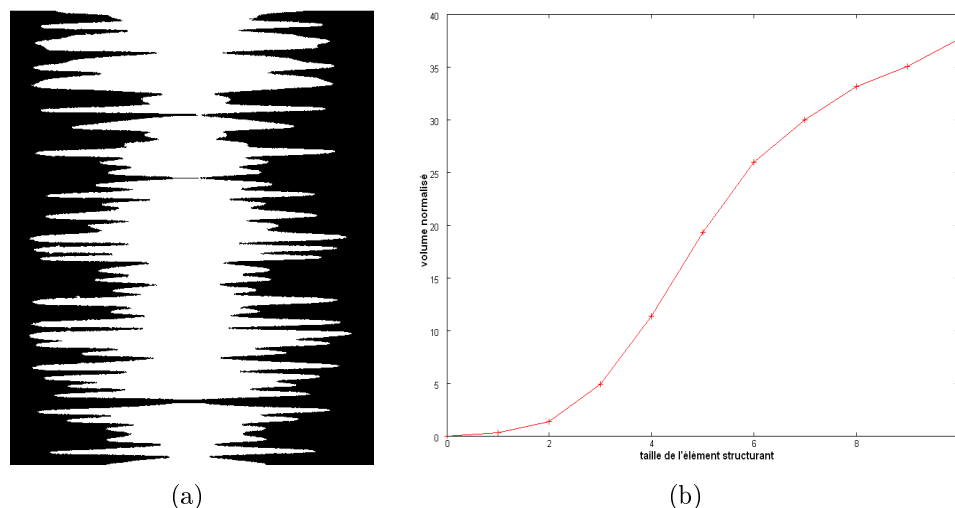


FIGURE 4.7 – (a) Image sous-exposée segmentée. (b) Courbe des granulométries des ouvertures.

Remarquons que la courbe est «quasi-plate» pour les premières valeurs d'éléments structurants. Ceci signifie que l'image sous-exposée est presque invariante par rapport à l'ouverture (ce qui a tendance à confirmer notre hypothèse). Dès que l'on dépasse la taille de l'élément structurant qui correspond à la dilatation modélisant au mieux la sous-exposition de l'image en question, on observe un «décollage» de la courbe. Enfin notons que, le fait que la différence entre l'image sous-exposée et son ouverture ne soit pas exactement nulle, s'explique par le fait que la dilatation n'est pas tout à fait égale à une sous-exposition, mais constitue cependant une bonne approximation.

En pratique, pour détecter la sous-exposition en utilisant cette méthode, on définit l'indicateur *InvarianceOuverture* qui est égal au volume normalisé de la différence entre l'image et son ouverture par un élément structurant de type carré de taille 2×2 . Cet élément structurant est très fin et donc bien adapté aux images contenant des hautes fréquences. Dans le cas de notre résolution, cette taille d'élément structurant est suffisante même pour des degrés importants de sous-exposition. Une valeur de *InvarianceOuverture* proche de 0 signifierait que l'image est presque invariante par rapport à l'ouverture, et donc qu'elle a probablement subi une sous-exposition.

D'une manière similaire, l'effet de la sur-exposition est très proche de celui d'une érosion, elle peut donc être détectée en vérifiant l'invariance par rapport à la fermeture.

4. Nous appelons volume normalisé, la somme des niveaux de gris de l'image divisée par son nombre de pixels

On définit pour ce faire, l'indicateur *InvarianceFermeture* comme étant le volume normalisé de la différence entre l'image et sa fermeture par le même élément structurant défini précédemment. Le comportement de *InvarianceFermeture* est similaire à celui de *InvarianceOuverture* (avec la différence qu'il détecte la sur-exposition).

Enfin, notons que l'invariance (par rapport à la fermeture et l'ouverture) est toujours vérifiée pour les images ne contenant que des basses fréquences. Cette méthode est à appliquer seulement aux images à hautes fréquences (l'effet de la mauvaise exposition étant, de toutes façons, négligeable sur les basses fréquences).

4.4.2 Dissymétrie des pics et des vallées

Une autre catégorie d'approches de détection a été développée. Elle est basée sur le fait que dans les pistes son optiques mal exposées, les pics et les vallées ont des caractéristiques différentes et ne sont en général pas symétriques (pics pointus et vallées creuses ou vice-versa). Cette dissymétrie peut être calculée de plusieurs manières. Nous avons proposé deux méthodes.

Au niveau des surfaces

La dissymétrie due à une mauvaise exposition fait que la surface des pics est différente de celle des vallées. Les opérateurs morphologiques constituent des outils efficaces pour mesurer ces deux surfaces. En effet, une fermeture avec un élément structurant vertical supprime les pics de la piste son, et une ouverture avec le même élément structurant remplit ses vallées. Ainsi, la surface des pics peut être calculée comme étant le volume de l'image de différence entre l'image originale et sa fermeture avec un élément structurant vertical. De même, la surface des vallées correspond au volume de l'image de différence entre l'image originale et son ouverture avec le même élément structurant.

En fonction de la taille choisie de l'élément structurant vertical, cette méthode permet de mesurer, soit une partie, soit la totalité de la surface des pics ou des vallées (figure 4.8).

Remarquons que pour une «petite» taille d'élément structurant (inférieure ou égale à 3 pixels dans le cas de notre résolution), ces valeurs correspondent à la surface des *sommets* des pics et celle des *fonds* des vallées. Dans le cas d'une piste sous-exposée, les pics étant très pointus, une fermeture avec un élément structurant vertical de «petite» taille éliminera une grande partie des pics (qui correspond à la surface des pics dont la largeur est inférieure ou égale à l'élément structurant utilisé). Si l'on applique une ouverture avec le même élément structurant à cette image, cette fermeture ne remplira qu'une «petite» partie des vallées, puisque celles-ci sont creuses et beaucoup moins pointues. Ainsi, la différence entre la surface des pics (éliminés par la fermeture) et celle des vallées (remplies

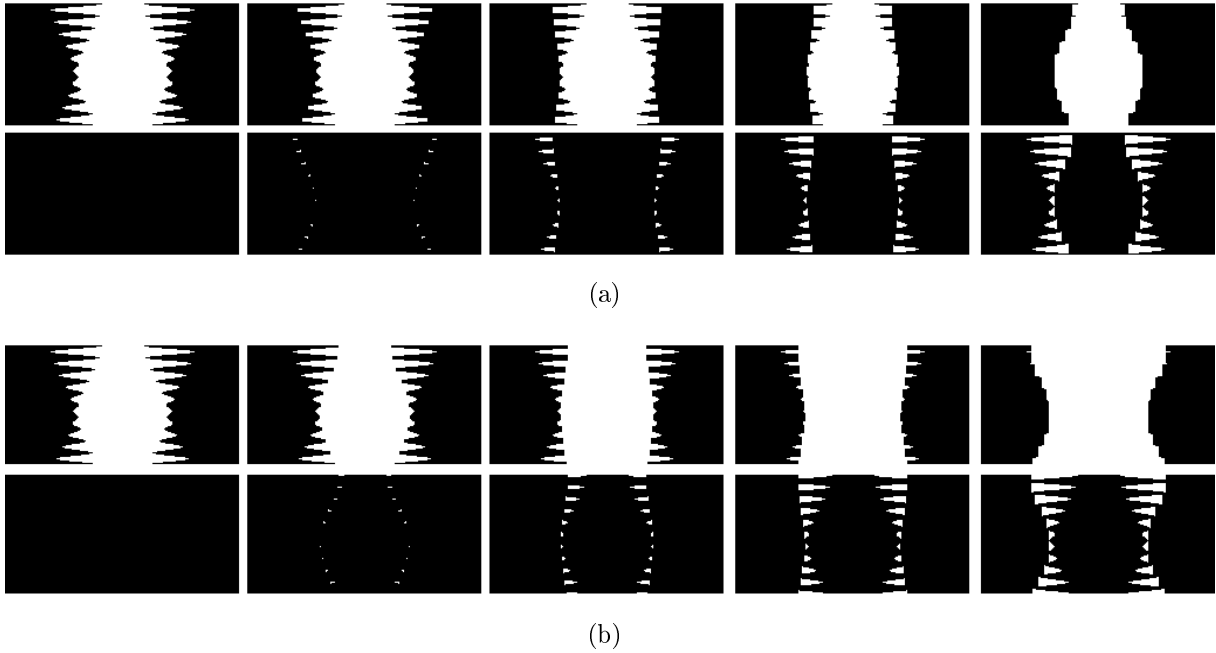


FIGURE 4.8 – (a) Succession d'ouvertures avec des éléments structurants verticaux de tailles croissantes et les différences correspondantes (entre l'image originale et ses ouvertures). (b) Succession de fermetures avec des éléments structurants verticaux de tailles croissantes et les différences correspondantes (entre l'image originale et ses fermetures).

par l'ouverture) nous informe sur la bonne ou la mauvaise exposition de la piste son optique. Cette différence est positive (surface des pics supérieure à celle des vallées) dans le cas d'une piste sous-exposée, négative dans le cas d'une piste sur-exposée et quasi-nulle dans le cas d'une exposition correcte.

Il est intéressant de noter qu'avec des tailles plus grandes d'éléments structurants, on ne compare plus la surface des sommets des pics avec celles des fonds des vallées, mais la totalité des surfaces des pics avec la totalité de celles des vallées. Si l'on procède ainsi, la différence mentionnée précédemment aura un comportement inverse (négative dans le cas sur-exposé et positive dans le cas sous-exposé). Pour comprendre ce comportement, prenons encore une fois l'exemple d'une piste sous-exposée (pics pointus et vallées creuses, figure 4.9(a)). Dans le cas d'une petite taille d'élément structurant, la fermeture supprime toute la partie pointue des pics (figure 4.9(b)) et l'ouverture ne remplit qu'une partie négligeable des vallées puisque celles-ci ne sont pas pointues (figure 4.9(c)). Mais pour des grandes tailles d'éléments structurants, ces opérateurs morphologiques supprimeront la totalité des pics et des vallées, et puisque les pics sont beaucoup plus pointus, leur surface sera beaucoup moins importante (figures 4.9(d) et 4.9(e)).

Notons finalement que la taille de l'élément structurant permettant de mesurer la totalité de la surface des pics et des vallées, dépend des fréquences présentes dans chaque

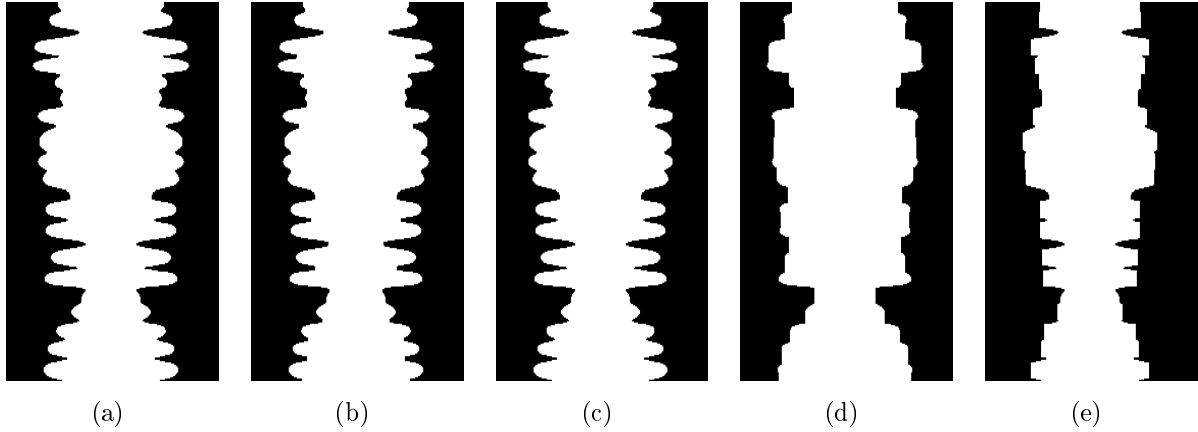


FIGURE 4.9 – (a) Piste son optique sous-exposée. (b) Fermeture par un élément structurant de taille 3. (c) Ouverture par un ES de taille 3. (d) Fermeture par un ES de taille 13. (e) Ouverture par un ES de taille 13.

image et n'est donc pas facile à déterminer. C'est la raison pour laquelle nous avons choisi dans cette méthode de travailler avec des petites tailles d'éléments structurants.

Pour mettre en pratique cette méthode, nous procédons comme suit :

Soit I une image d'une piste son optique, B un élément structurant vertical de taille 3, γ_B et φ_B respectivement l'ouverture et la fermeture par l'élément structurant B . On calcule le volume $v1$ de l'image de différence entre I et $\gamma_B(I)$, et le volume $v2$ de l'image de différence entre $\varphi_B(I)$ et I . $v1$ correspond alors à la surface des vallées, $v2$ à celles des pics. On définit l'indicateur de mauvaise exposition *DissymetrieSurfaces* par *DissymetrieSurfaces* = $v2 - v1$. Cet indicateur sera alors positif dans le cas d'une sur-exposition, négatif dans le cas d'une sous-exposition et quasi-nul dans le cas d'une exposition correcte.

Au niveau de la courbure

Pour calculer la courbure d'un pic ou d'une vallée, nous utilisons la méthode proposée dans la thèse de R. Effciaud pour le calcul des contraintes locales dans la propagation des lignes de partage des eaux [26], que nous adaptons à notre cas comme suit :

Soit p un pixel faisant partie du contour de l'image segmentée. Nous calculons dans un voisinage de p d'une taille déterminée t , le nombre $n1$ des pixels blancs et le nombre $n2$ des pixels noirs (figure 4.10).

Il est intuitif que la différence $n_1 - n_2$ augmente avec la courbure locale du contour. Pour étudier plus en détail la relation entre cette différence et la courbure (ou le rayon de courbure), considérons un disque D de rayon r et une fenêtre carrée de longueur t centrée

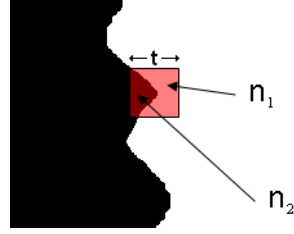
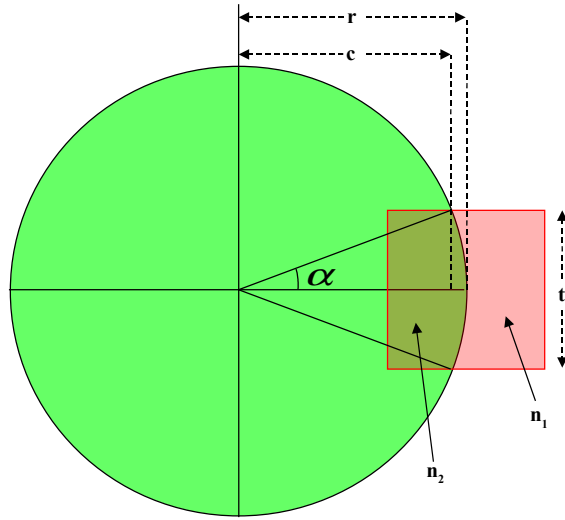


FIGURE 4.10 – Calcul de la courbure.

sur un point de contour de D (figure 4.11).

FIGURE 4.11 – n_1 et n_2 sur un disque de rayon r .

Après quelques calculs géométriques, nous obtenons :

$$n_2 = r^2 \cdot \alpha - c \cdot \frac{t}{2} + \left(\frac{t}{2} - (r - c) \right) \cdot t \text{ et}$$

$$n_1 = t^2 - n_2.$$

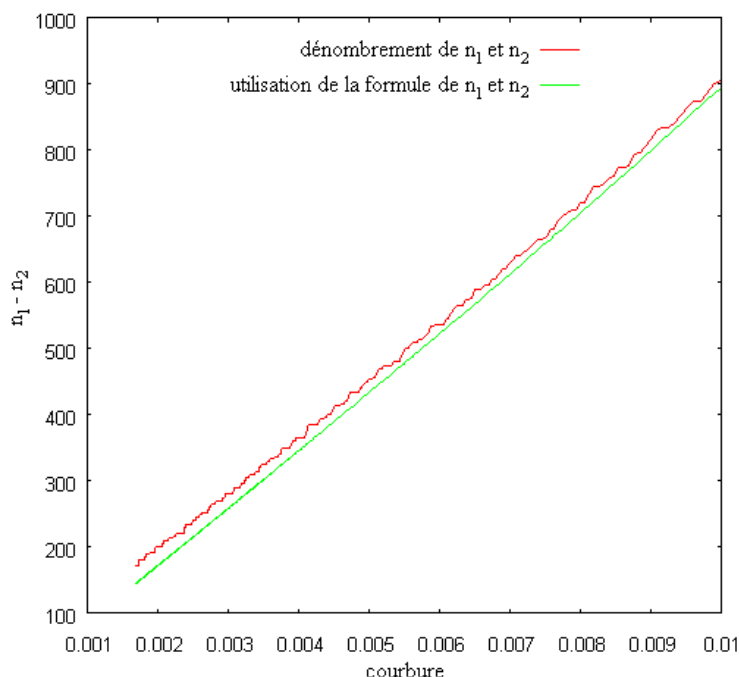
où :

$$c = \sqrt{r^2 - \left(\frac{t}{2} \right)^2} \text{ et}$$

$$\alpha = \arctan \left(\frac{t}{2c} \right).$$

La courbe de la figure 4.12 montre l'allure de la différence $n_2 - n_1$ en fonction de la courbure $1/r$ pour une valeur de t égale à 50 pixels. Notons que sur cette figure, le calcul est effectué soit en utilisant les formules de n_1 et n_2 précédemment mentionnées, soit en dénombrant le nombre de pixels n_1 et n_2 dans un disque discret de rayon r . Les différences que l'on peut observer entre les deux méthodes de calcul sont dues à la discrétisation.

En ce qui concerne les pistes son, la différence $n_1 - n_2$ est positive dans le cas d'un pic et négative dans le cas d'une vallée. De même, plus cette différence est importante, plus le pic (ou la vallée) est pointu.

FIGURE 4.12 – $n_1 - n_2$ en fonction de la courbure $1/r$.

Pour exploiter cette propriété, nous dénombrons :

- le nombre np des pixels du contour pour lesquels $n_1 \geq 2 \cdot n_2$ (ou pixels des pics) et
- le nombre nv des pixels du contour pour lesquels $n_2 \geq 2 \cdot n_1$ (ou pixels des vallées).

Puisque les pics sont plus pointus que les vallées dans une image sous-exposée, np sera supérieur à nv . L'inverse est vrai pour une image sur-exposée.

En pratique, nous définissons l'indicateur *DissymetrieCourbures* comme étant le rapport $DissymetrieCourbures = \frac{np-nv}{np+nv}$. La taille t du voisinage est fixée à 5 pixels. Comme le montre la figure 4.13, cette valeur est quasiment nulle si l'image est correctement exposée, positive si elle est sous-exposée et négative si elle est sur-exposée.

4.4.3 Etude comparative des différents indicateurs

Nous avons donc développé les quatre indicateurs suivants :

- *InvarianceOuverture* : Invariance par rapport à l'ouverture : détecte seulement la sous-exposition ;
- *InvarianceFermeture* : Invariance par rapport à la fermeture : détecte seulement la sur-exposition ;
- *DissymetrieSurfaces* : Dissymétrie au niveau des surfaces : détecte la sur et la sous-exposition ;

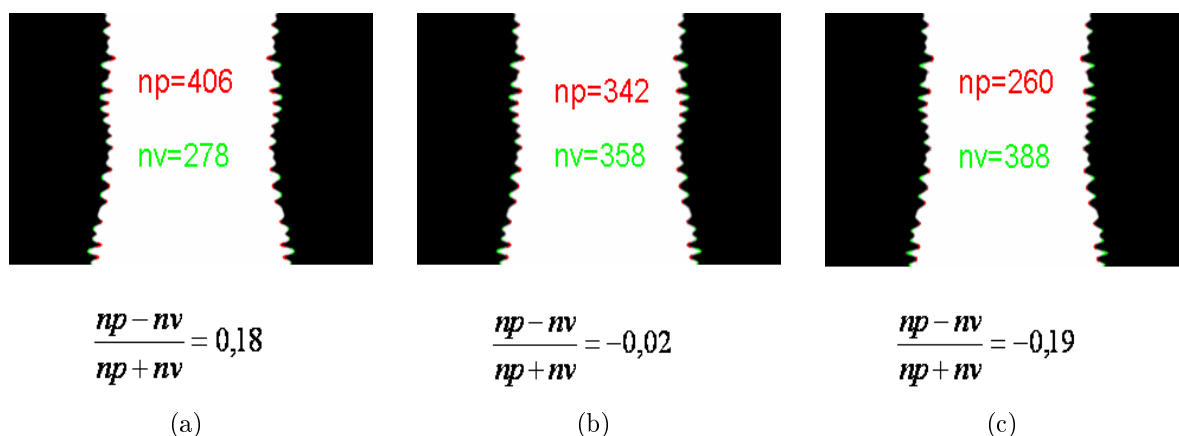


FIGURE 4.13 – Rapport entre pixels de pics et pixels de vallées dans le cas d’une image sous-exposée (a) , correctement exposée (b) , et sur-exposée (c) .

- *DissymetrieCourbures* : Dissymétrie au niveau des courbures : détecte la sur et la sous-exposition.

Nous avons calculé ces indicateurs sur quelques images des pistes pour lesquelles nous avons simulé une mauvaise exposition (figure 4.14), sur les pistes de cross-modulation (figure 4.15) et sur les images prises de la piste son optique de la bande annonce de «Délice Paloma» (figure 4.16).

Notons que pour tous ces nuages de points, les abscisses représentent les valeurs de chaque indicateur et les ordonnées représentent l’indice de l’image dans la piste son optique. Rappelons que les points ayant la même ordonnée correspondent à une même partie de la piste son obtenue avec différentes expositions (normale, sur-exposée et sous-exposée).

Ces indicateurs ont été calculés sur des images contenant des hautes fréquences. Pour des raisons pratiques, ils ont été calculés image par image. Leur application sur une succession de plusieurs images peut être considérée à condition que cette succession contienne des hautes fréquences. Nous verrons dans la section suivante, comment ces indicateurs vont être utilisés en pratique.

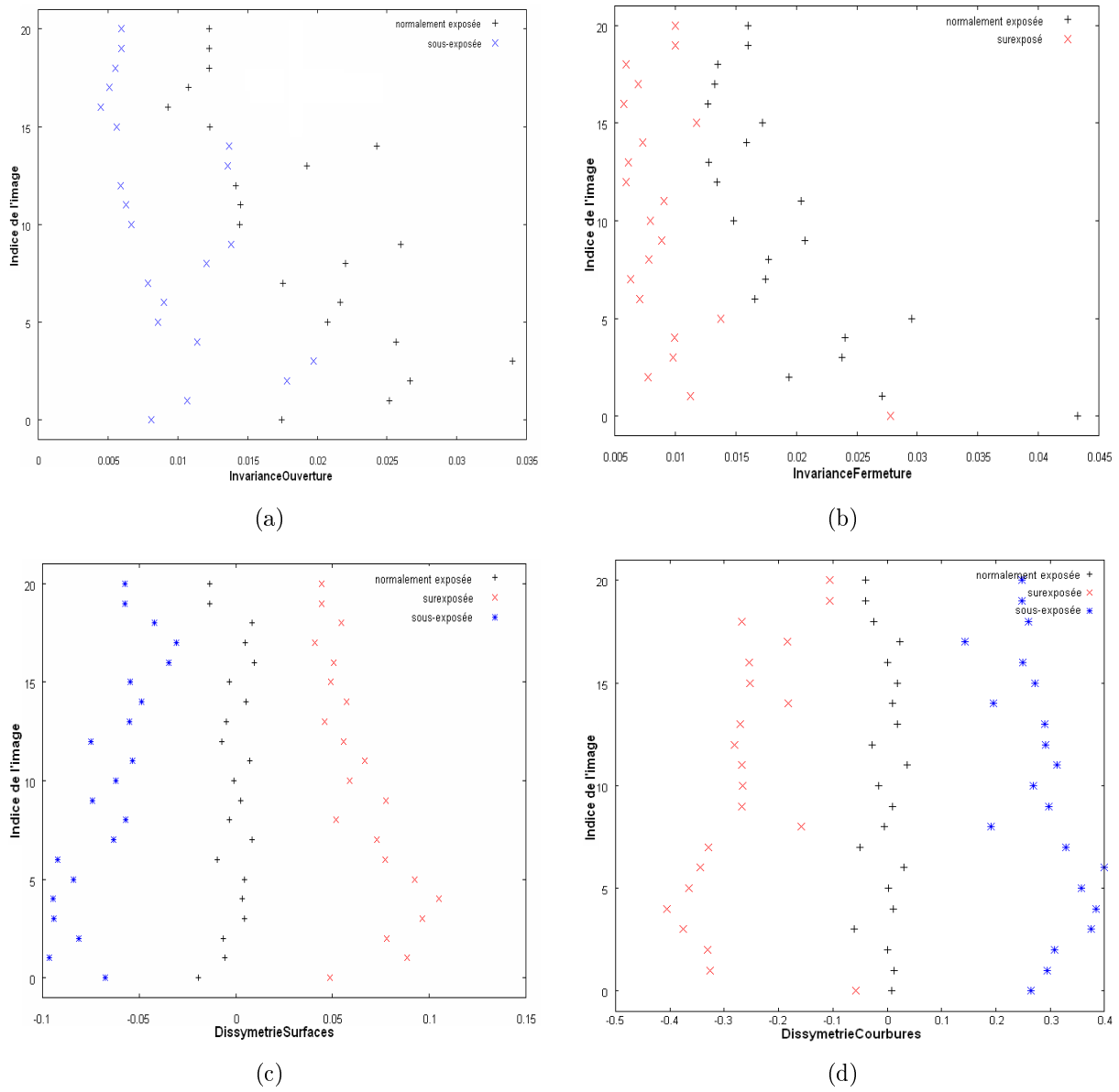


FIGURE 4.14 – Nuages de points donnant, pour des pistes son ayant subi une sur/sous-exposition simulée, la valeur des indicateurs pour des images avec différentes expositions.

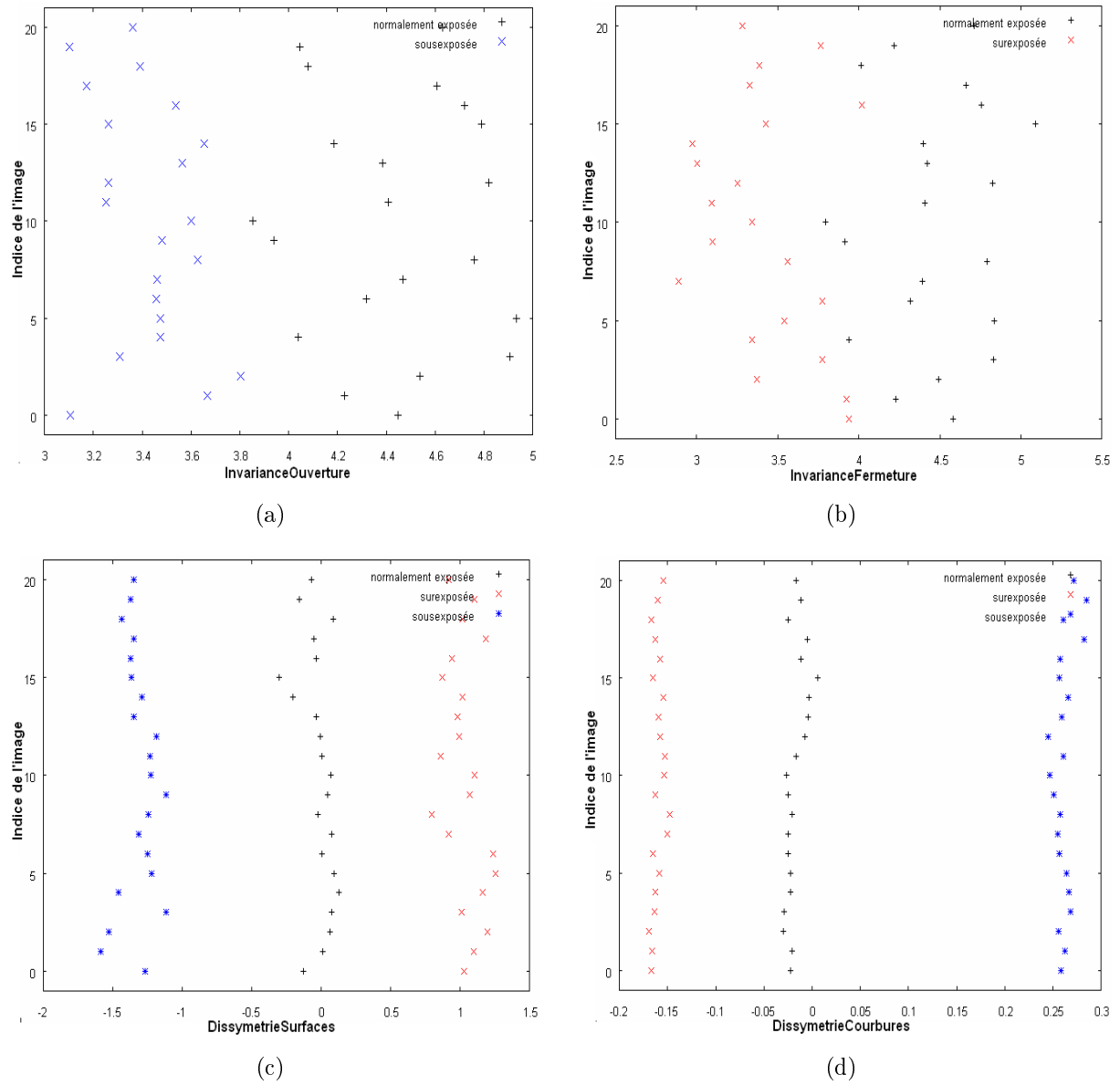


FIGURE 4.15 – Nuages de points donnant, pour des pistes son de cross-modulation, la valeur des indicateurs pour des images avec différentes expositions.

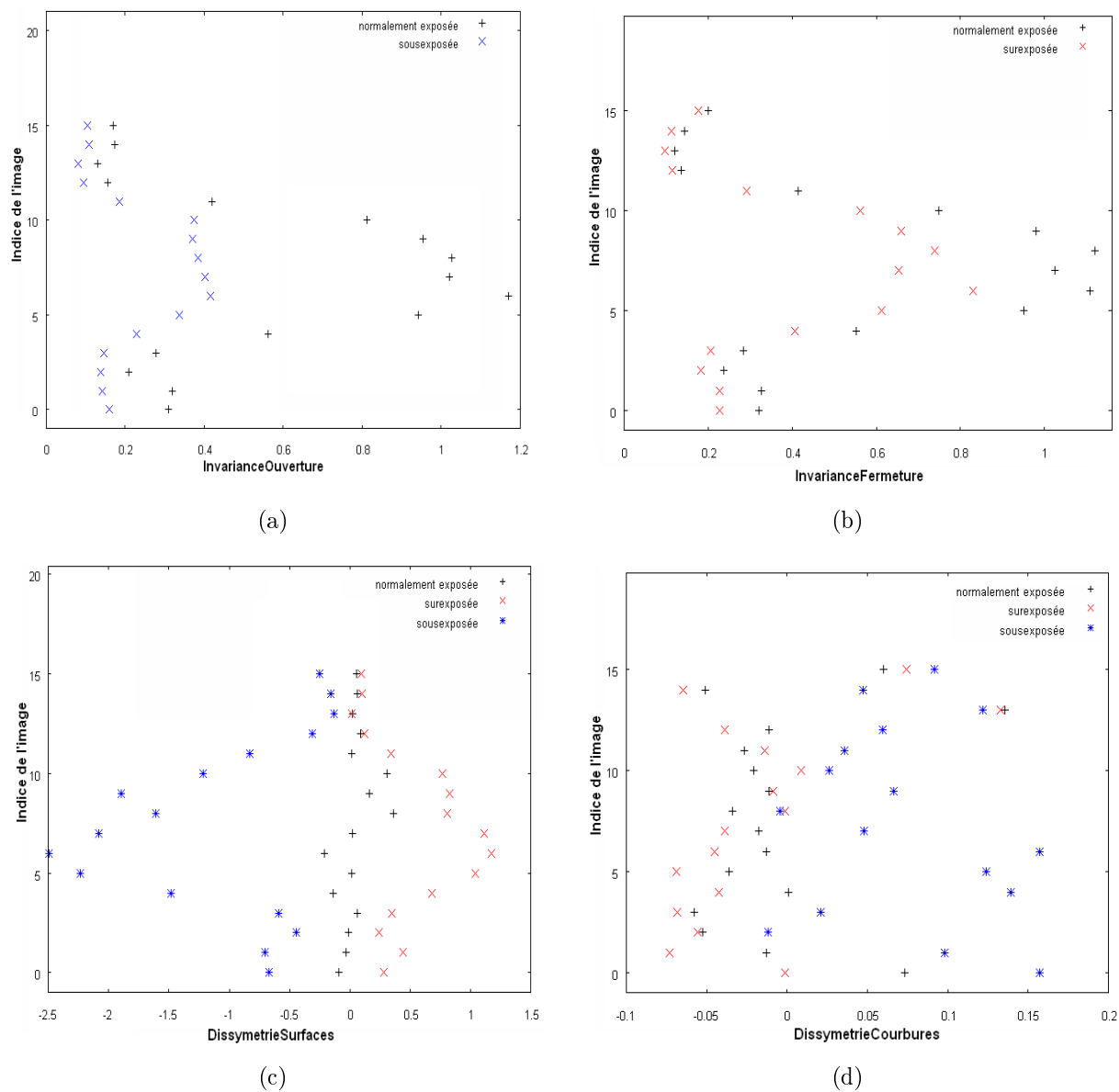


FIGURE 4.16 – Nuages de points donnant, pour la piste son de la bande annonce de «Délice Paloma», la valeur des indicateurs pour des images avec différentes expositions.

Remarquons que :

- La valeur de *InvarianceOuverture* (respectivement *InvarianceFermeture*) est plus proche de 0 dans le cas d'une image sous-exposée (respectivement sur-exposée) que dans le cas de l'image correspondante correctement exposée.
- Aucun des indicateurs développés pour le moment ne permet de placer un seuil au dessus (ou au dessous) duquel l'image est sur ou sous-exposée. Ceci ne met pas en cause l'utilité de ces indicateurs. Ils pourraient en effet servir dans un système de restauration itératif, dans lequel on calcule la valeur de l'indicateur à chaque itération et l'on teste s'il évolue dans le sens de l'exposition correcte.
- Remarquons que, le degré de mauvaise exposition étant bien marqué dans le cas des pistes simulées, les indicateurs arrivent à bien distinguer les images correspondant aux différentes expositions.
- Cette distinction est encore plus marquée dans le cas des pistes de cross-modulation. Ceci est dû, d'une part à la forte présence des hautes fréquences dans ces pistes (qui se traduit par une importante dissymétrie dans le cas d'une mauvaise exposition) et à la périodicité de ce type de pistes d'autre part (qui implique la multiplication des écarts entre les valeurs des indicateurs).
- *DissymetrieCourbures* ne distingue pas les différentes expositions pour le cas des images de «Délice Paloma». Ceci est dû au fait que l'on est à la limite de la résolution. La figure 4.17 montre le même résultat avec une résolution trois fois plus importante.
- Enfin, notons que les variations des indicateurs entre les images successives sont dues aux variations des fréquences présentes dans chaque image. En effet, certaines images contiennent beaucoup plus de hautes fréquences que d'autres.

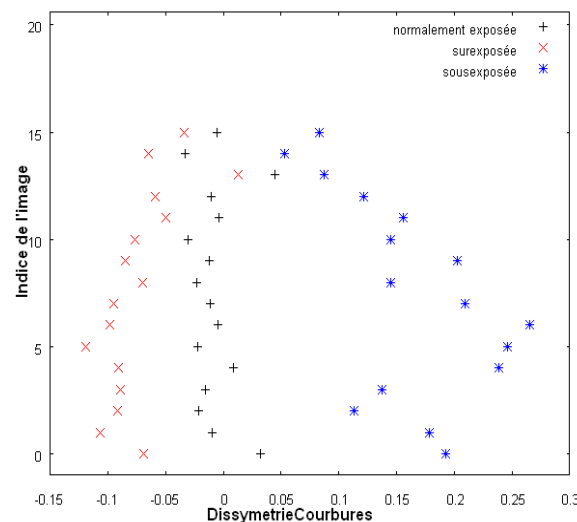


FIGURE 4.17 – Nuages de points donnant la valeur de l'indicateur de dissymétrie des courbures en fonction de l'indice de l'image dans le cas de la bande annonce de «Délice Paloma» avec une résolution trois fois plus importante.

4.4.4 Stabilité des indicateurs par rapport au seuil

Nous avons vu que certains indicateurs nécessitent une segmentation préalable de l'image, soit parce qu'ils ne peuvent être calculés que sur une image binaire (cas de la méthode développée pour le calcul de courbure), soit pour réduire l'effet du bruit. Malheureusement, tous les indicateurs ne réagissent pas par rapport au seuil de la même façon. En d'autres termes, un seuil «mal choisi» pourrait par exemple donner à une image normalement exposée, le comportement d'une image sur/sous-exposée (vis-à-vis des indicateurs) ou vice versa.

Pour illustrer ces propos, la figure 4.18 montre la valeur des indicateurs en fonction du seuil, pour une piste son prise avec trois expositions différentes de la bande annonce de «Délice Palome». Remarquons que pour l'image en question, les indicateurs *DissymetrieSurfaces* et *DissymetrieCourbures* ont un comportement normal (*DissymetrieSurfaces* négatif dans le cas sous-exposé, positif dans le cas sur-exposé et quasi-nul dans le cas correctement exposée et vice-versa pour *DissymetrieCourbures*) pour des valeurs de seuil allant de 80 à 120.

Nous concluons qu'il est important de déterminer la valeur «correcte» du seuil avant l'application des différents indicateurs. Cette valeur dépend des conditions d'acquisition de chaque piste son, elle a été empiriquement fixée à 100 pour toutes les images sur lesquelles nous avons travaillé. Nous pensons cependant, qu'une approche permettant de rendre les indicateurs plus robustes vis-à-vis de la valeur du seuil mérite d'être étudiée.

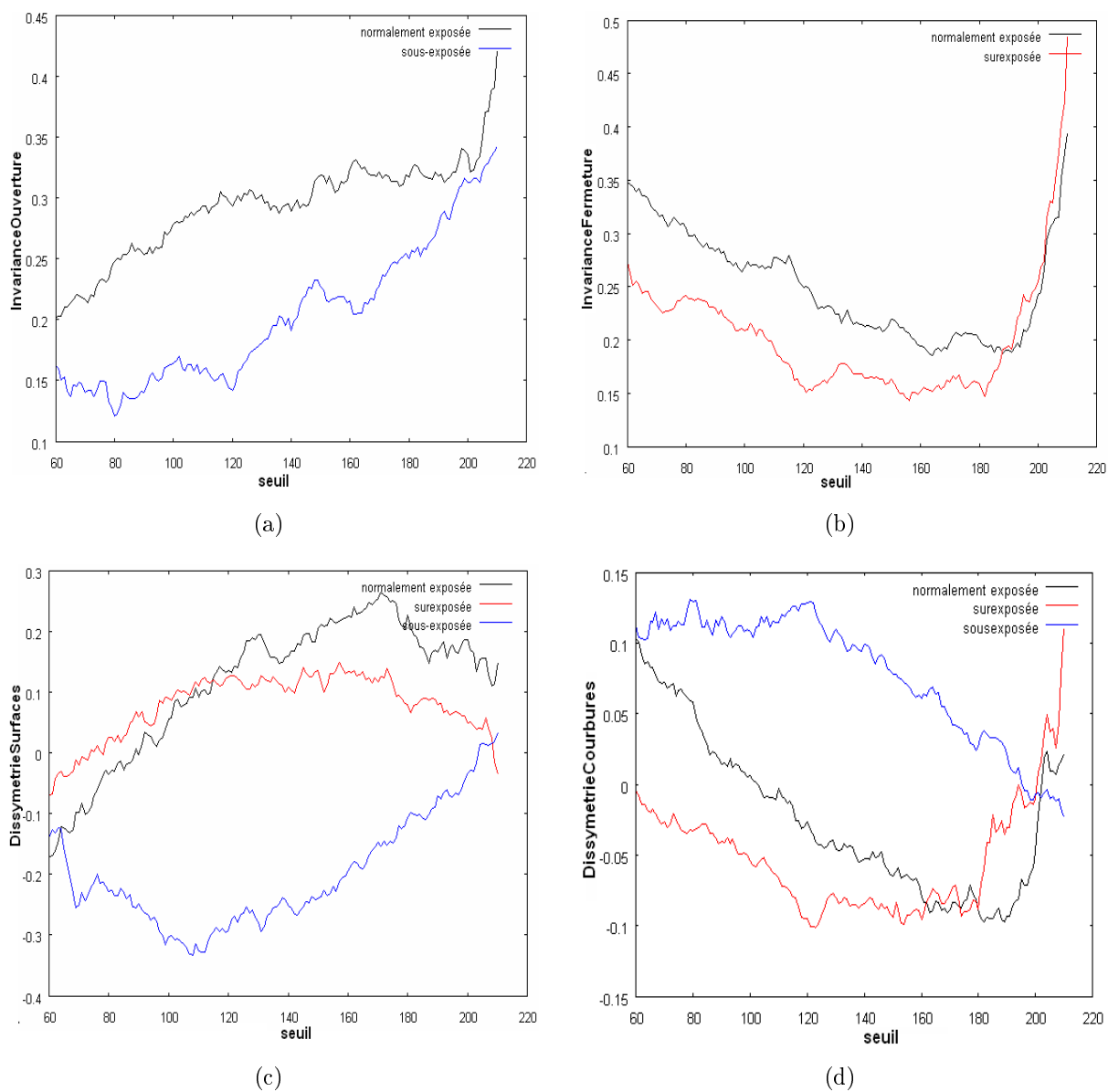


FIGURE 4.18 – Valeurs des indicateurs en fonction du seuil pour une image prise de «Délice Paloma».

4.5 Correction de la mauvaise exposition

4.5.1 Application d'une érosion (ou d'une dilatation)

Nous avons vu précédemment que la sous-exposition est très proche d'une dilatation morphologique avec un certain élément structurant. L'érosion étant l'opération duale de la dilatation, elle constituerait donc un bon candidat pour la restauration d'une telle piste son. La restauration d'une piste sur-exposée peut être réalisée d'une manière similaire en appliquant une dilatation.

A titre d'illustration, nous avons généré numériquement la piste son optique correspondant à un «S», sur laquelle nous avons simulé une sous-exposition. Enfin, nous avons appliqué à la piste sous-exposée une érosion avec un élément structurant carré de taille 2×2 . Notons qu'un élément structurant plus large remplirait l'espace entre les pics, ce qui n'est pas souhaité.

D'un point de vue audio, nous avons remarqué que le «S» sous-exposé contient un souffle que l'on n'aperçoit plus après l'application de l'érosion. La figure 4.19 montre les spectrogrammes de ces trois signaux audio. Remarquons que la sous-exposition introduit des hautes fréquences qui n'étaient pas présentes dans le signal audio original. Ces hautes fréquences sont moins marquées (niveaux de gris moins sombres) dans le signal audio correspondant à la piste son érodée.

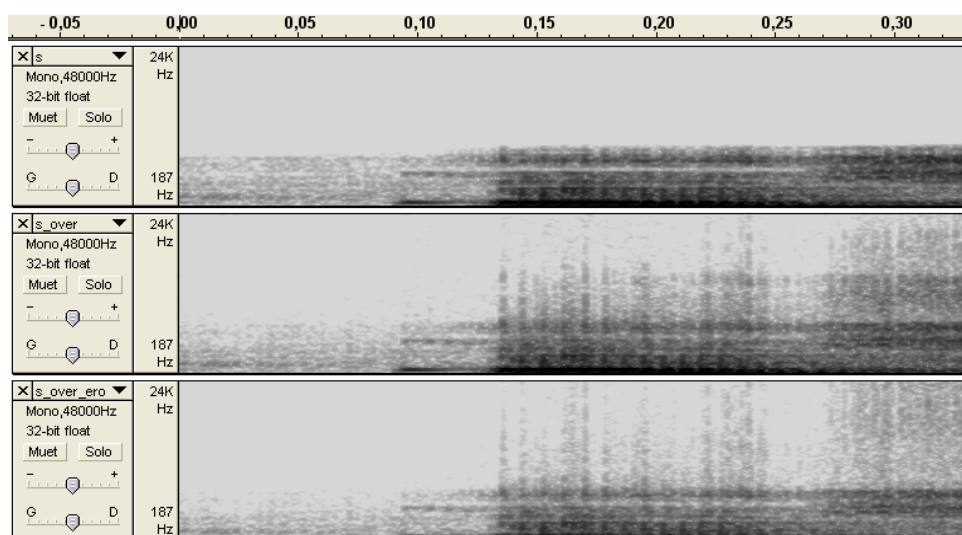


FIGURE 4.19 – Spectrogramme correspondant à un «S», de haut en bas : normalement exposé, sous-exposé, puis sous-exposé sur lequel on a appliqué une érosion.

La taille de l'élément structurant «idéal» avec lequel il faudrait éroder l'image peut être déterminée en utilisant les différents indicateurs développés. Cependant, une érosion avec un élément structurant de taille unitaire est largement suffisante, voire trop importante

dans certains cas. Pour cette raison, il nous a semblé important d'introduire la notion d'«éléments structurants de taille réelle» lors du calcul de l'érosion.

La notion de «morphologie mathématique à échelle continue» (ou continuous-scale morphology) n'est pas nouvelle. En effet, depuis 1992 plusieurs équipes de recherche ont développé des méthodes de cette discipline à base d'équations aux dérivées partielles ([3, 14, 71]). Parallèlement à ces travaux, d'autres développements concernant l'évolution des courbes en utilisant les ensembles de niveaux ont eu lieu [56]. D'autres travaux ont montré que ces opérateurs morphologiques peuvent aussi être obtenus à partir de transformées en distance [10]. F. Guichard et al. ont montré dans une étude la relation entre toutes ces méthodes [29].

Dans ce travail, nous avons développé une méthode plus simple par rapport à ce que l'on a trouvé dans la littérature : pour effectuer l'érosion d'une image I en utilisant un élément structurant de taille réelle, nous augmentons la résolution de I par une méthode d'interpolation bilinéaire (figure 4.20(b)). Nous calculons ensuite l'érosion de l'image interpolée. Cette érosion est seulement calculée pour les pixels de l'image interpolée qui «correspondent» à des pixels dans l'image originale (figure 4.20(c)). Les valeurs de ces pixels sont reportées sur l'image de sortie (figure 4.20(d)). Notons que si la résolution a été augmentée par un facteur de n et que l'érosion dans l'espace interpolé a été effectuée avec un élément structurant de taille t , nous parlerons d'un élément structurant de taille t/n .

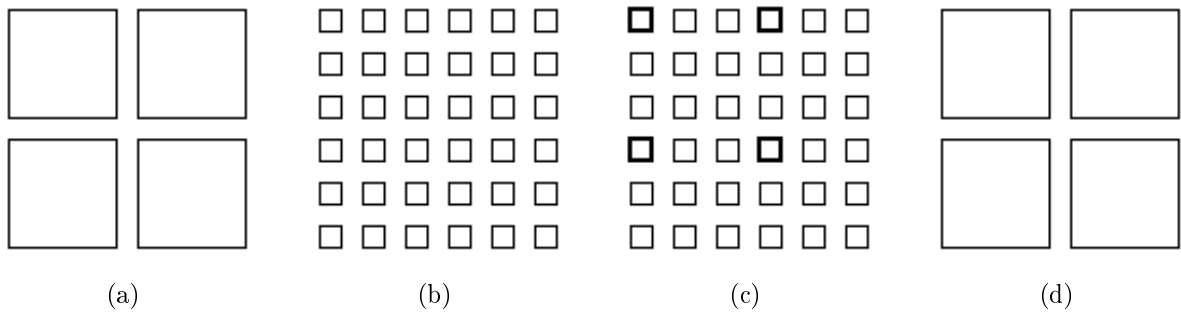


FIGURE 4.20 – Schéma de l'érosion proposée : Image originale (a). Image interpolée (b). Mise en évidence des pixels pour lesquels l'érosion est calculée (c). Image de sortie (d).

Pour rappel, l'interpolation bilinéaire d'un pixel $I(x', y')$ est calculée en fonction des valeurs de ses pixels voisins (figure 4.21) en utilisant les formules :

$$I(x', y') = (y_2 - y') \cdot I(x', y_1) + (y' - y_1) \cdot I(x', y_2)$$

avec $I(x', y_1) = (x_2 - x') \cdot I(x_1, y_1) + (x' - x_1) \cdot I(x_2, y_1)$,
et $I(x', y_2) = (x_2 - x') \cdot I(x_1, y_2) + (x' - x_1) \cdot I(x_2, y_2)$.

A titre d'illustration, une succession d'érosions morphologiques avec des éléments structurants ronds de tailles croissantes en utilisant une résolution dix fois plus impor-

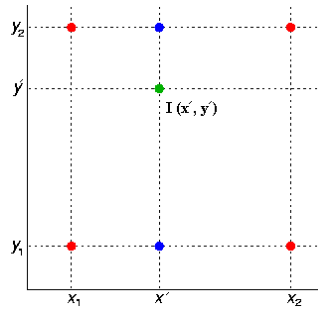


FIGURE 4.21 – Voisins d'un pixel dans l'image interpolée.

tante est montrée sur quelques formes géométriques dans la figure 4.22 et sur une piste de cross modulation dans la figure 4.23.

Pour détecter automatiquement la taille de l'élément structurant avec lequel il faudrait éroder (respectivement dilater) l'image pour corriger la sous-exposition (respectivement sur-exposition), on calcule pour chaque taille d'élément structurant, les valeurs des indicateurs pour l'image érodée (ou dilatée) correspondante. Nous avons choisi d'utiliser les indicateurs *DissymetrieSurfaces* et *DissymetrieCourbures* parce qu'ils permettent de savoir à partir de quelle taille d'élément structurant, on passe d'une sur-exposition à une sous-exposition (ou inversement).

La courbe de la figure 4.24(b) montre la valeur des indicateurs en fonction de la taille de l'élément structurant rond utilisé pour l'érosion de l'image sous-exposée de la figure 4.24(a). Remarquons que la taille optimale de l'élément structurant est de 0,6 puisqu'elle donne les valeurs de *DissymetrieSurfaces* et *DissymetrieCourbures* les plus proches de 0. Le résultat de l'érosion de l'image sous-exposée en question avec un élément structurant rond de cette taille est montré dans la figure 4.24(c).

De manière similaire, la courbe de la figure 4.25(b) montre la valeur de ces deux indicateurs sur la granulométrie des dilations morphologiques avec des éléments structuraux ronds de tailles croissantes pour l'image sur-exposée de la figure 4.25(a). La taille optimale de l'élément structurant est dans ce cas de 0,2 et le résultat de la restauration est montré dans la figure 4.25(c).

Nous avons remarqué que la taille optimale des éléments structuraux ne varie que très peu le long d'une même piste son, il suffit donc de la déterminer une seule fois. Rappelons cependant que cette valeur doit être calculée sur une partie de la piste son qui contient des hautes fréquences, puisque les indicateurs développés ne sont pertinents que pour ce type d'images.

Bien que les résultats de cette méthode soient généralement satisfaisants, elle reste très consommatrice en termes de temps de calcul. Nous avons, pour cette raison, proposé une autre méthode de restauration.

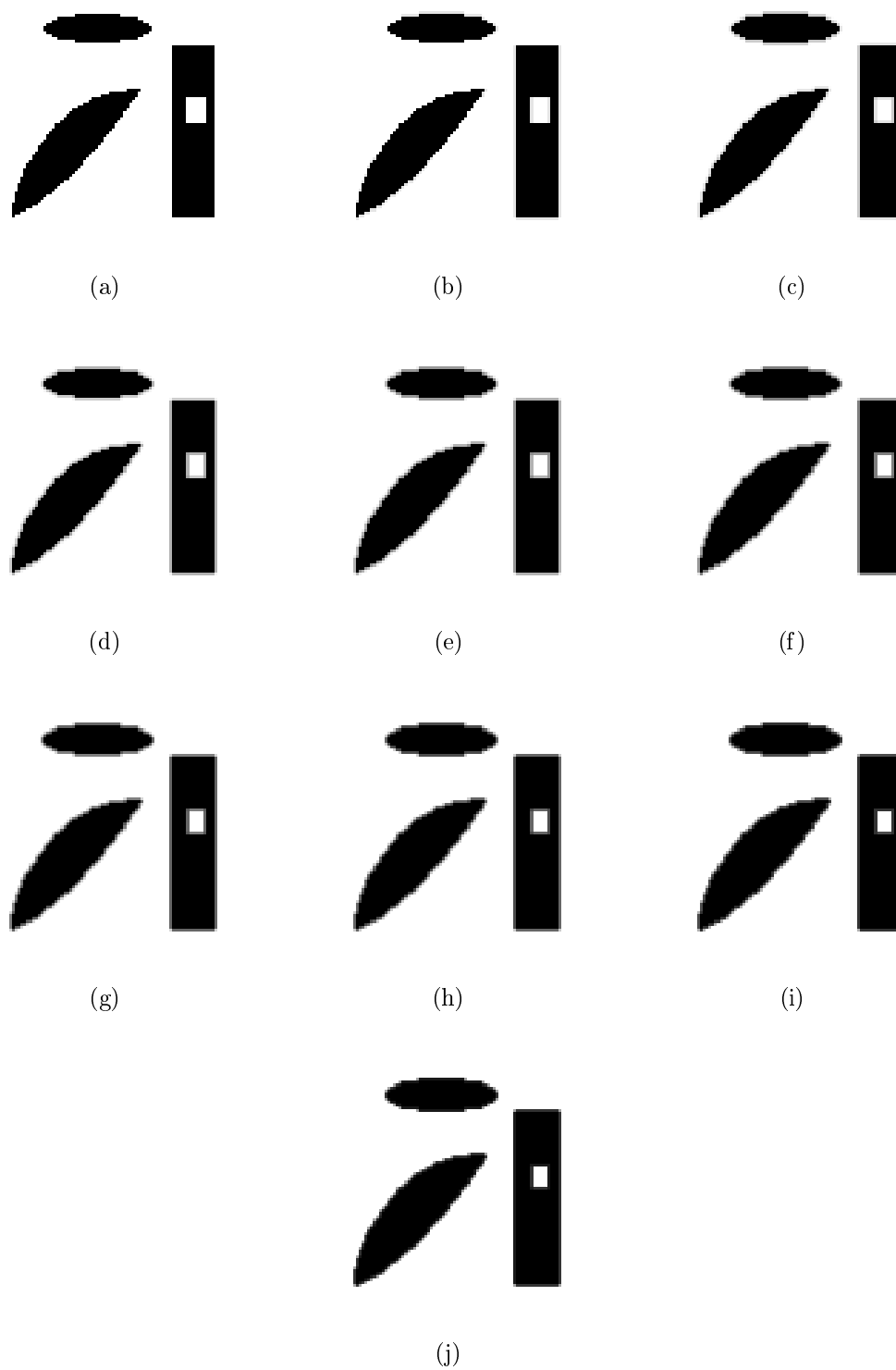


FIGURE 4.22 – Erosion de formes géométriques avec un élément structurant rond de taille 0 (a), 0,1 (b), 0,2 (c), 0,3 (d), 0,4 (e), 0,5 (f), 0,6 (g), 0,7 (h), 0,8 (i) et 0,9 (j).

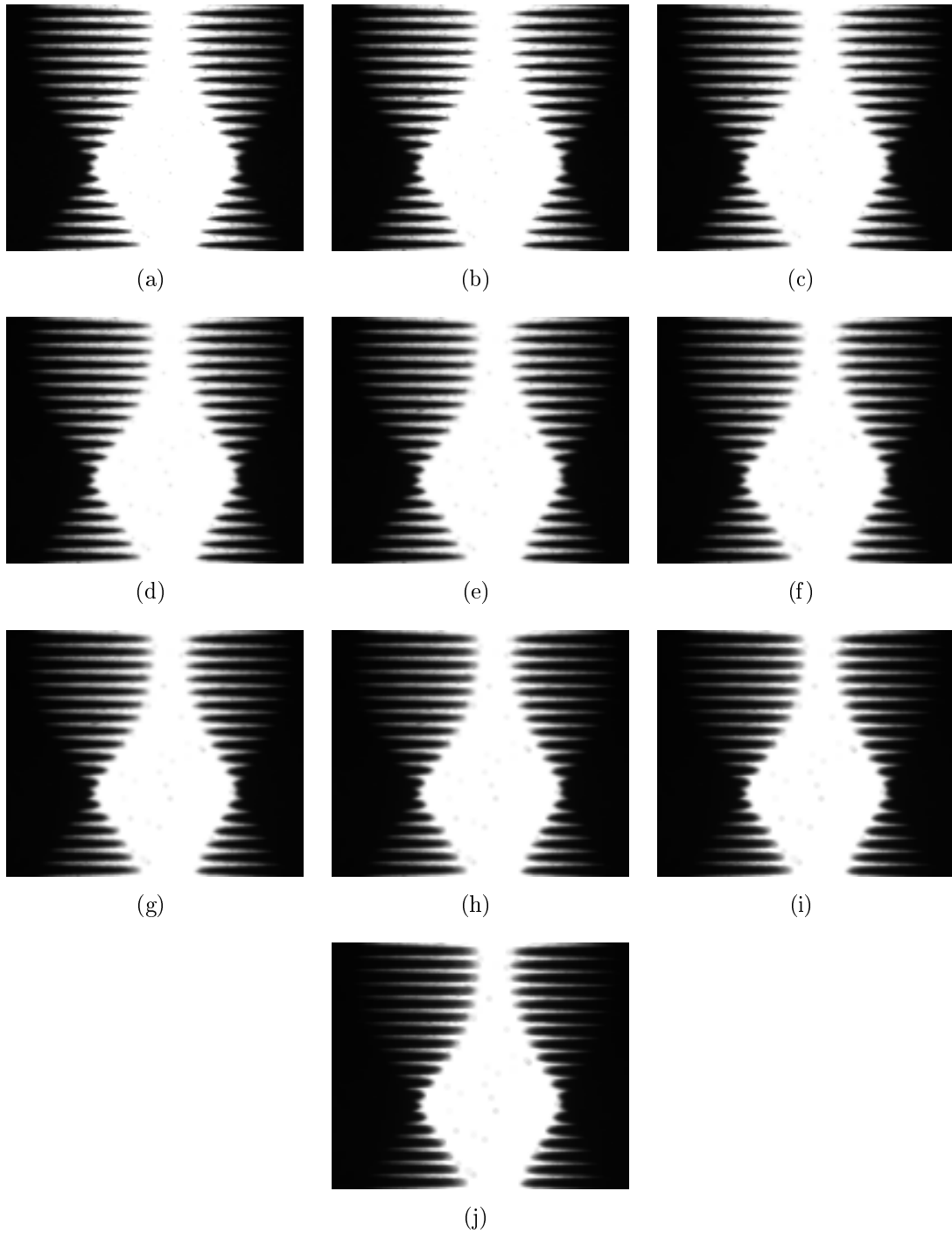


FIGURE 4.23 – Erosion d’une piste de cross-modulation avec un élément structurant rond de taille 0 (a), 0,1 (b), 0,2 (c), 0,3 (d), 0,4 (e), 0,5 (f), 0,6 (g), 0,7 (h), 0,8 (i) et 0,9 (j).

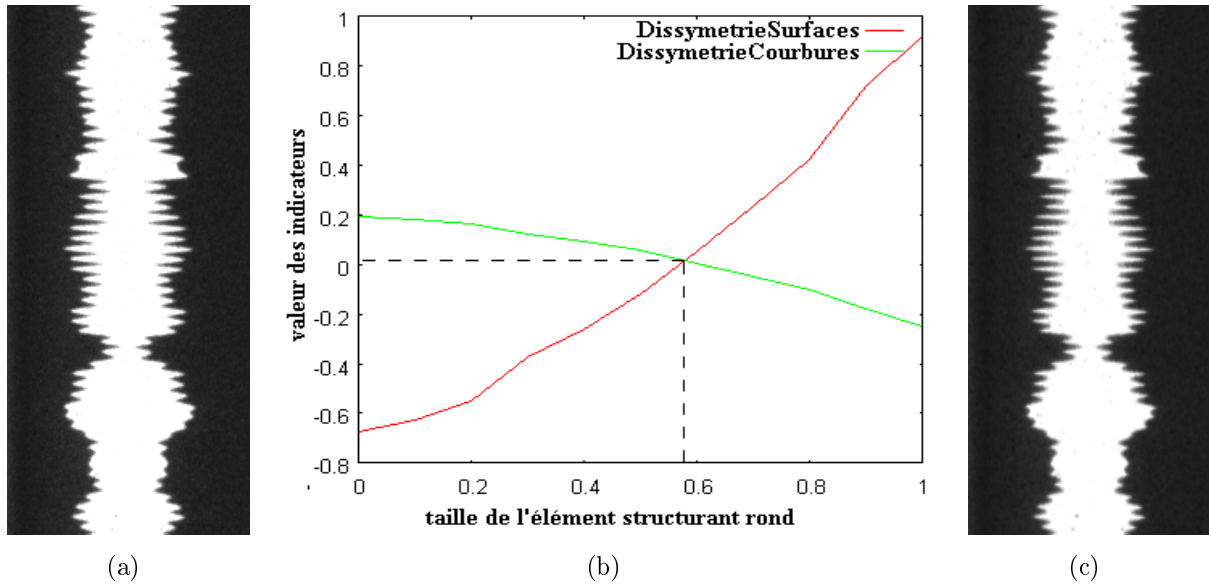


FIGURE 4.24 – (a) Image sous-exposée. (b) Valeurs des indicateurs en fonction de la taille de l'élément structurant. (c) Image érodée en utilisant la valeur optimale de l'élément structurant.

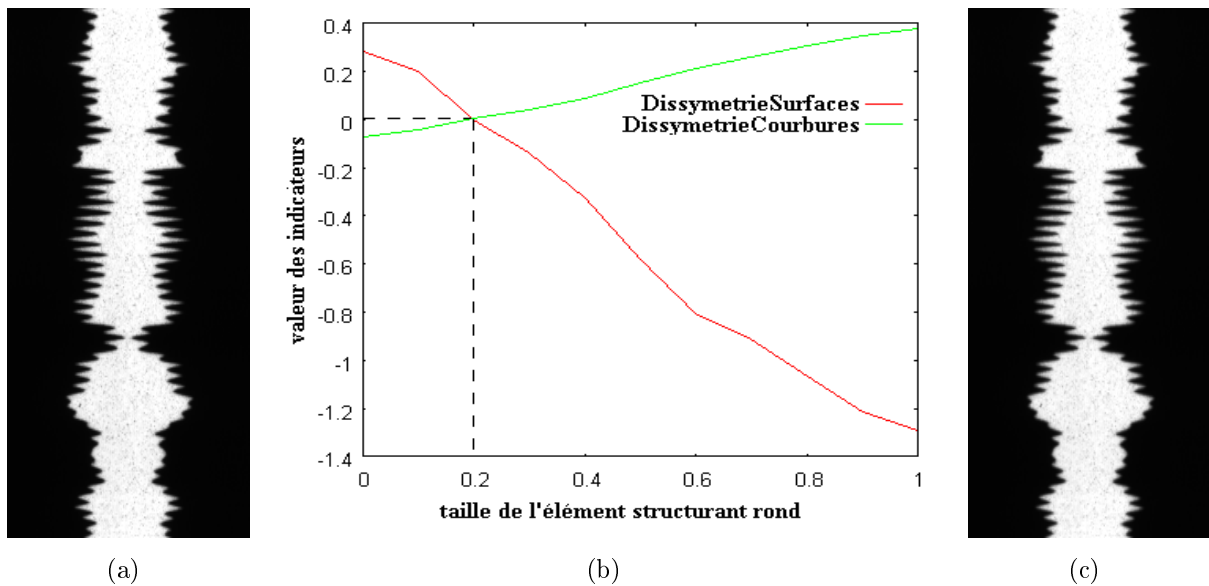


FIGURE 4.25 – (a) Image sur-exposée. (b) Valeurs des indicateurs en fonction de la taille de l'élément structurant. (c) Image dilatée en utilisant la valeur optimale de l'élément structurant.

4.5.2 L'application d'une table de correspondance (LUT) sur les niveaux de gris

Nous utilisons dans cette partie, une variante de la méthode proposée par Bhukhanwala et Ramabadran [9].

La fonction de transfert du film est simulée par $\phi(x) = \frac{255}{1+e^{-c \cdot x}}$ où c est une constante contrôlant le contraste, sa valeur par défaut est 0,75.

La fonction $push(p1, \delta x) = \phi(\phi^{-1}(p1) + \delta x)$ permet de corriger l'exposition de δx . Une valeur positive de δx permet de corriger la sur-exposition, et une valeur négative de δx permet de corriger la sous-exposition. La figure 4.26 montre l'allure de la fonction ϕ ainsi que celle de $push$ en fonction des différentes valeurs de δx .

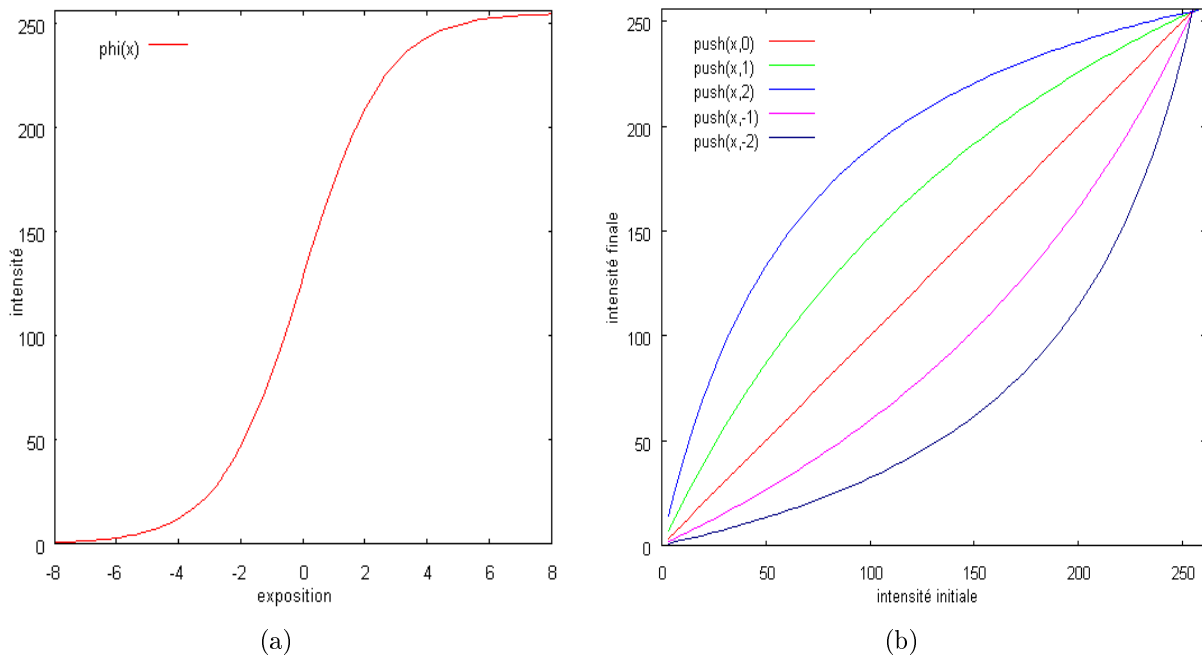


FIGURE 4.26 – (a) Fonction H&D. (b) Evolution de la courbe $push$ en fonction des différentes valeurs de δx .

La correction sera ensuite apportée en appliquant aux intensités de l'image, la courbe $push$ avec la «bonne» valeur de δx . Contrairement aux méthodes trouvées dans la littérature, cette valeur ne sera pas déterminée de manière à mettre le plus en évidence les détails (ce qui risquerait dans notre cas d'augmenter le bruit), mais de telle sorte que l'on obtienne des valeurs optimales pour les indicateurs développés. Nous avons choisi d'utiliser les indicateurs *DissymetrieSurfaces* et *DissymetrieCourbures* pour la même raison évoquée précédemment.

Pour illustrer cette méthode, nous avons appliqué la courbe $push$ avec des valeurs de

δx allant de 0 à 3 sur l'image sur-exposée de la figure 4.27(a). La courbe de la figure 4.27(b) montre l'évolution des indicateurs *DissymetrieSurfaces* et *DissymetrieCourbures* en fonction de δx . Enfin, la figure 4.27(c) montre l'image obtenue avec la valeur optimale de δx , valeur à partir de laquelle *DissymetrieSurfaces* passe du positif (qui représente les pistes sur-exposées) au négatif (qui représente les pistes sous-exposées) et vice-versa pour *DissymetrieCourbures*.

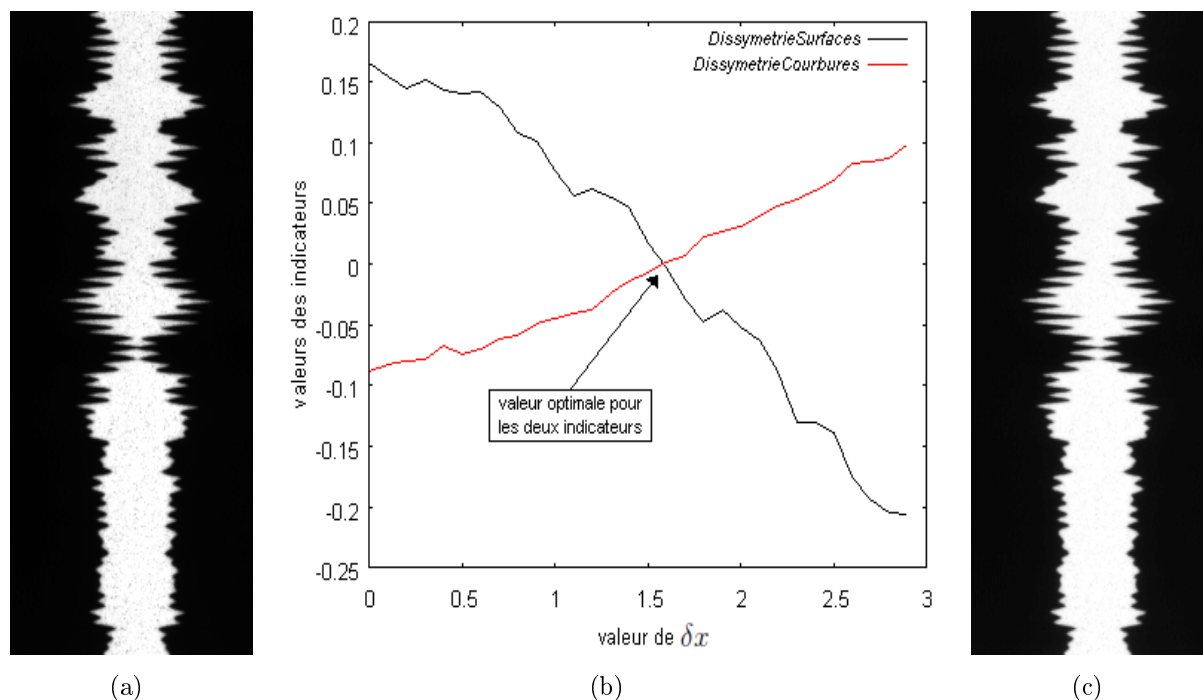


FIGURE 4.27 – (a) Image sur-exposée. (b) Valeurs des indicateurs en fonction de δx . (c) Image restaurée en utilisant la valeur optimale de δx .

4.5.3 Analyse des deux méthodes

Nous avons testé ces deux méthodes sur l'ensemble de nos bases de données et nous avons fait les remarques suivantes :

L'érosion (ou la dilatation) permet d'annuler le souffle dû à la mauvaise exposition, mais elle ne permet cependant pas de revenir à l'image originale. En effet, pour une reproduction exacte de l'image originale (correctement exposée), il faudrait que l'érosion soit effectuée avec un élément structurant de taille variable. La figure 4.28 montre une piste sous-exposée, la piste correspondante normalement exposée et l'image donnant, pour chaque pixel, la taille de l'élément structurant rond avec lequel il faudrait éroder l'image sous-exposée pour l'obtention de l'image correctement exposée. Remarquons qu'au niveau

des pics, la taille de l'élément structurant idéal est importante. En se dirigeant vers les vallées, cette taille a tendance à diminuer. Une méthode automatique permettant de calculer la taille de l'élément structurant qu'il faudrait appliquer pour l'érosion (ou la dilatation) de chaque pixel fait partie de nos perspectives.

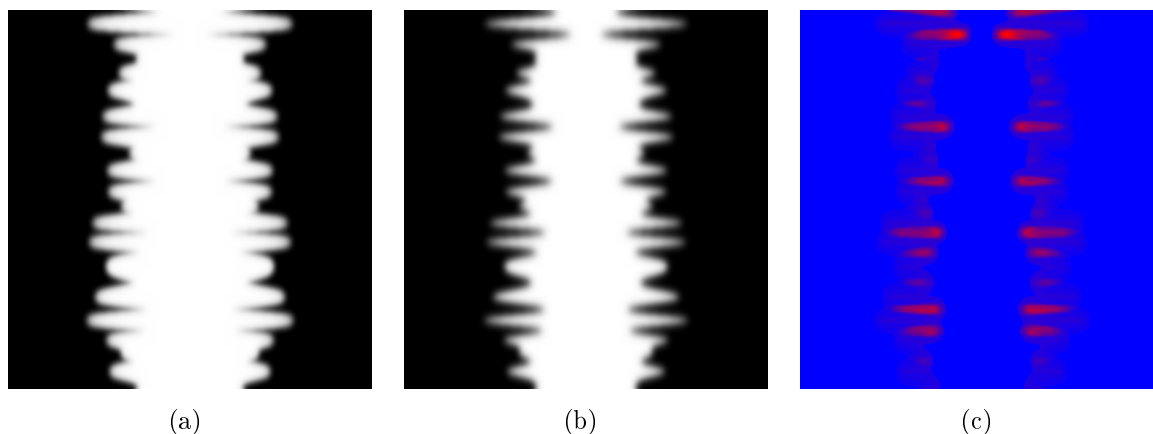


FIGURE 4.28 – (a) Image sous-exposée. (b) Image correctement exposée. (c) Image donnant pour chaque pixel, la taille de l'élément structurant rond avec lequel il faudrait éroder l'image (a) pour obtenir l'image (b), le rouge représente une taille importante et le bleu une taille nulle.

La deuxième méthode, quant à elle, apporte une correction qui est plus «naturelle» puisqu'elle travaille directement sur les intensités lumineuses. Cependant, dans certains cas où le degré de mauvaise exposition est important, les zones de transitions (entre la région claire et les régions sombres) sont quasi-saturées et une correction de ce type ne peut reconstruire complètement ces zones de transition. De même, pour certaines images qui nécessitent une correction importante, le niveau de gris de la région sombre de l'image peut être altéré considérablement en utilisant cette méthode (fig. 4.29). Enfin, une correction importante de ce type se traduit aussi par une réduction du nombre de niveaux de gris de l'image. Ceci a une influence sur la dynamique du signal audio et se traduit par un bruit de fond au niveau sonore.

L'application d'une table de correspondance constitue donc une bonne correction de l'exposition des pistes son si le degré de mauvaise exposition n'est pas très important. Cette méthode a l'avantage d'être très efficace et peut être calculée en temps réel. Dans le cas d'une mauvaise exposition importante, l'approche morphologique semble plus appropriée. Celle-ci est cependant plus coûteuse en termes de temps de calcul, notamment pour des grandes tailles d'éléments structurants. Notons que la combinaison des deux méthodes est une piste de recherche qui mérite d'être exploitée.

Plus de détails sur ces deux méthodes de restauration ainsi que sur leur intégration dans le système global de restauration seront donnés dans le chapitre suivant.

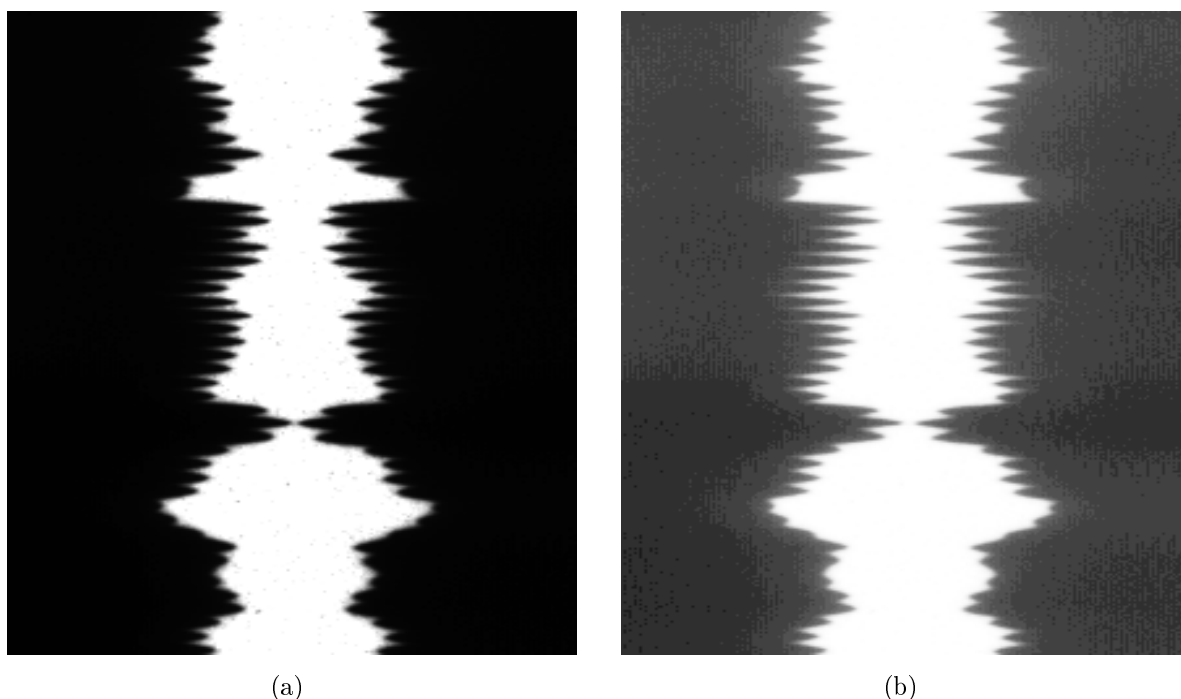


FIGURE 4.29 – (a) Image sur-exposée. (b) Image corrigée en appliquant la courbe *push*.

4.6 Tests d'écoute

La 7^{ème} et 8^{ème} comparaisons de chaque série de tests (voir section 1.7) correspondent à des pistes sous-exposée et sur-exposée avant et après restauration (en utilisant l'érosion ou la dilatation morphologique).

Les résultats de ces tests sont donnés dans le tableau 4.2.

Nombre de personnes trouvant	Piste sous-exposée	Piste sur-exposée
le son avant restauration de meilleure qualité	0	0
le son après restauration de meilleure qualité	1	1
qu'il n'y a pas de différence	11	5

TABLE 4.2 – Résultats des tests en aveugle pour les pistes mal exposées.

Malheureusement, ces pistes ne présentent pas un grand degré de mauvaise exposition et la plupart des personnes n'ont, en conséquence, pas trouvé de différence entre la version mal exposée et la version corrigée. Il est intéressant de noter que les seules deux personnes ayant trouvé une différence, sont deux participantes âgées de 24 et 27 ans.

Nous avons voulu aussi tester l'aptitude des participants à distinguer un son correctement exposé d'un son mal exposé. Ces deux signaux sonores correspondent respectivement

à *normalex480_13* et *sousex960_13* (voir tableau 4.1). Cette comparaison est la neuvième dans chaque série de tests. Les résultats de ce test sont donnés dans le tableau 4.3.

Nombre de personnes trouvant	
le son sous-exposé de meilleure qualité	0
le son correctement exposé de meilleure qualité	1
qu'il n'y a pas de différence	11

TABLE 4.3 – Comparaison par des tests en aveugle entre un son d'une piste correctement exposée et un son d'une piste sous-exposée.

La seule personne ayant préféré le son correspondant à la piste normalement exposée est la même participante de 27 ans mentionnée précédemment (celle de 24 ans n'ayant pas trouvé de différence ici).

Les archivistes du film trouvent cependant que le son correspondant à la piste correctement exposée ou corrigée contient moins de hautes fréquences que la piste mal exposée. Cette observation confirme le fait que la mauvaise exposition introduit des hautes fréquences que notre méthode de restauration a tendance à éliminer.

4.7 Conclusion

Nous avons mené une étude sur la détection et la restauration des pistes son optiques sur/sous-exposées au niveau image. Les indicateurs de détection développés sont basés, d'un côté sur les caractéristiques morphologiques des pistes son ayant subi une exposition incorrecte, et d'un autre sur la dissymétrie introduite par cette mauvaise exposition. Deux approches de correction ont été proposées : une approche morphologique qui permet de corriger directement la forme des pics et des vallées, et une seconde approche qui corrige les niveaux de gris de l'image en appliquant une table de correspondance (LUT) correspondant au degré de mauvaise exposition de la piste son.

Au niveau image, nous avons remarqué que les deux approches permettent de corriger la forme des pics et des vallées. L'approche morphologique donne de bons résultats même quand le degré de mauvaise exposition est important mais est plus consommatrice en termes de temps de calcul.

Les tests d'écoute en aveugle n'ont pas été concluants pour ce type de pistes, notamment à cause du fait que les participants n'arrivent pas à déceler ce genre de défauts. Cependant, les archivistes du film trouvent que notre méthode de restauration «filtre» les hautes fréquences introduites par la mauvaise exposition.

Plusieurs perspectives sont envisageables, notamment : l'amélioration des différents indicateurs vis-à-vis de leur robustesse par rapport au bruit ; le développement des méthodes de restauration morphologiques avec des éléments structurants de tailles variables ;

la combinaison de la méthode morphologique avec la méthode basée sur les tables de correspondances et l'étude de la possibilité d'effectuer les traitements proposés directement dans le domaine audio.

Enfin, tous ces développements sont à intégrer dans un système de restauration complet, qui corrige d'un côté les dégradations dues au vieillissement du film (poussières, moisissures, taches...etc) et d'un autre côté, les problèmes dus à la mauvaise exposition.

Système de restauration global

Dans les chapitres précédents, nous avons décrit les différents modules développés pour la restauration des pistes optiques à largeur variable et à densité variable, ainsi que des pistes présentant un problème de mauvaise exposition.

Plusieurs questions demeurent cependant posées :

- Dans le cas d’une piste présentant à la fois un problème de mauvaise exposition, et des défauts dus au vieillissement du film, laquelle des deux restaurations faudrait-il apporter en premier ? Le résultat final sera-t-il le même ?
- Comment est-ce que la communication est effectuée entre tous les modules développés ?
- Lesquels de ces modules sont parallélisables et lesquels ne le sont pas ?

Le but de ce chapitre est d’apporter une réponse à toutes ces questions.

5.1 Communication et ordonnancement

Vue globale des modules

La figure 5.1 montre l’organigramme de tous les modules développés.

Certains des modules développés servent à la paramétrisation du système et ne nécessitent d’être lancés qu’une fois au début du traitement. Ces modules comprennent :

- La détection du degré de déviation d’azimut (que ce soit pour les pistes VA ou VD) ;
- Le calcul de la valeur du «seuil» à utiliser pour la restauration des pistes VD (voir section 3.4.2) ;
- La détection du degré de mauvaise exposition.

Les autres modules développés sont consacrés à la restauration proprement dite.

Notons que la phase d’initialisation ne nécessite pas le même nombre d’images pour tous les traitements. Si elle ne nécessite qu’une seule image pour la détection du degré de mauvaise exposition, nous avons remarqué qu’en pratique, elle doit être exécutée sur succession de 10 images au moins pour avoir une mesure fiable de la valeur de déviation d’azimut, ainsi que la valeur du seuil permettant la restauration des pistes VD.

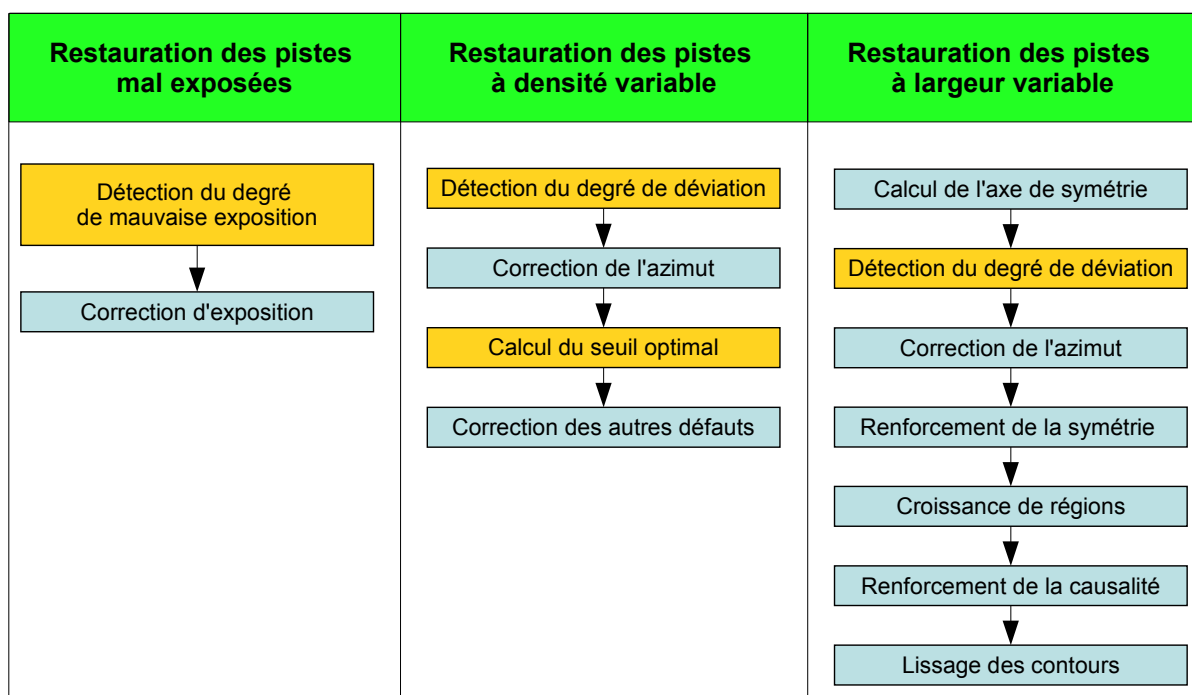


FIGURE 5.1 – Organigramme des modules développés. *Les modules en jaune servent à la paramétrisation du système.*

Dépendances et entrées/sorties

De manière plus détaillée, les figures 5.2, 5.3 et 5.4 montrent les dépendances qui existent entre tous les modules développés ainsi que les entrées et sorties de chacun de ces modules.

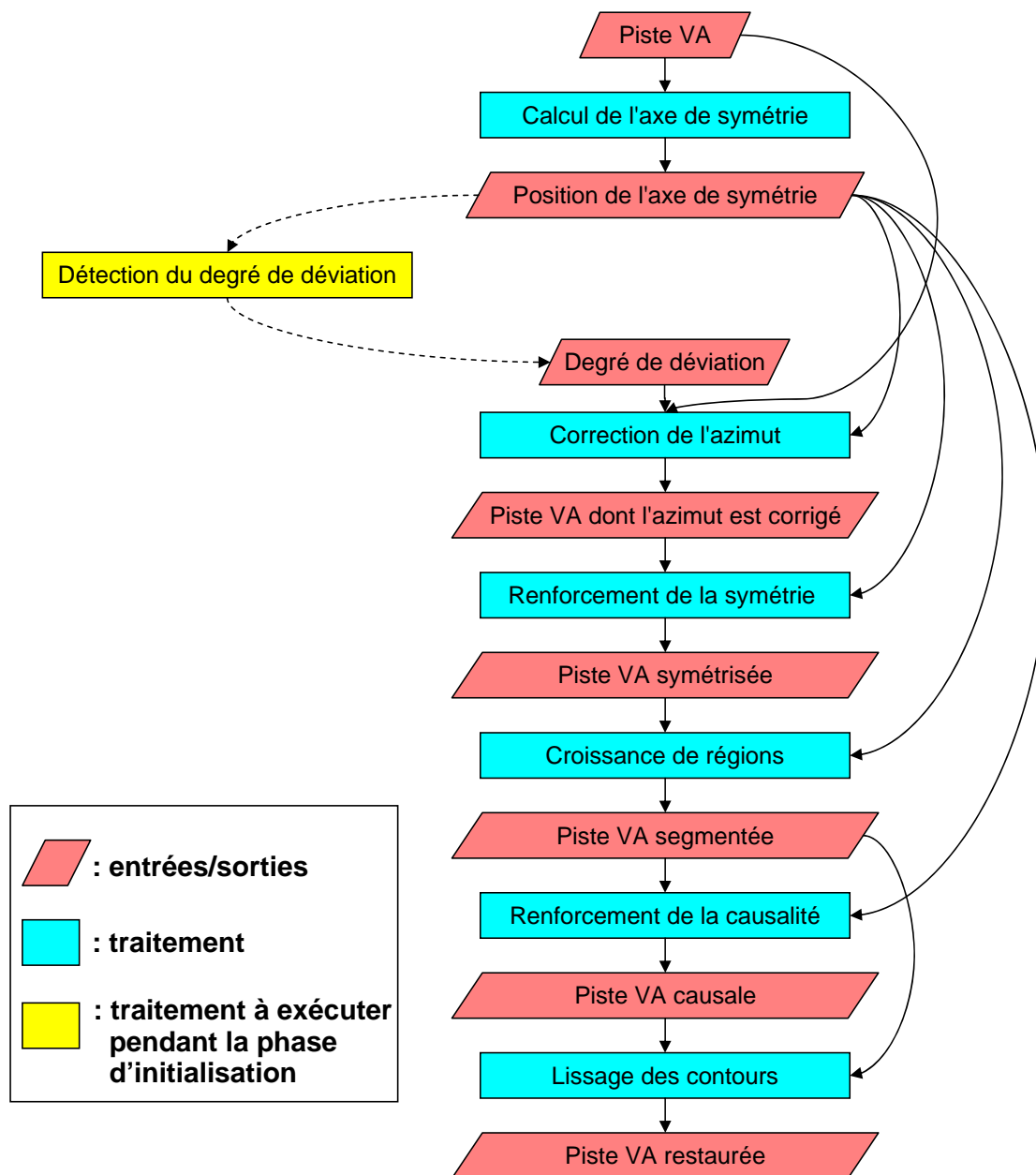


FIGURE 5.2 – Dépendances et entrées/sorties des modules de la restauration des pistes VA.

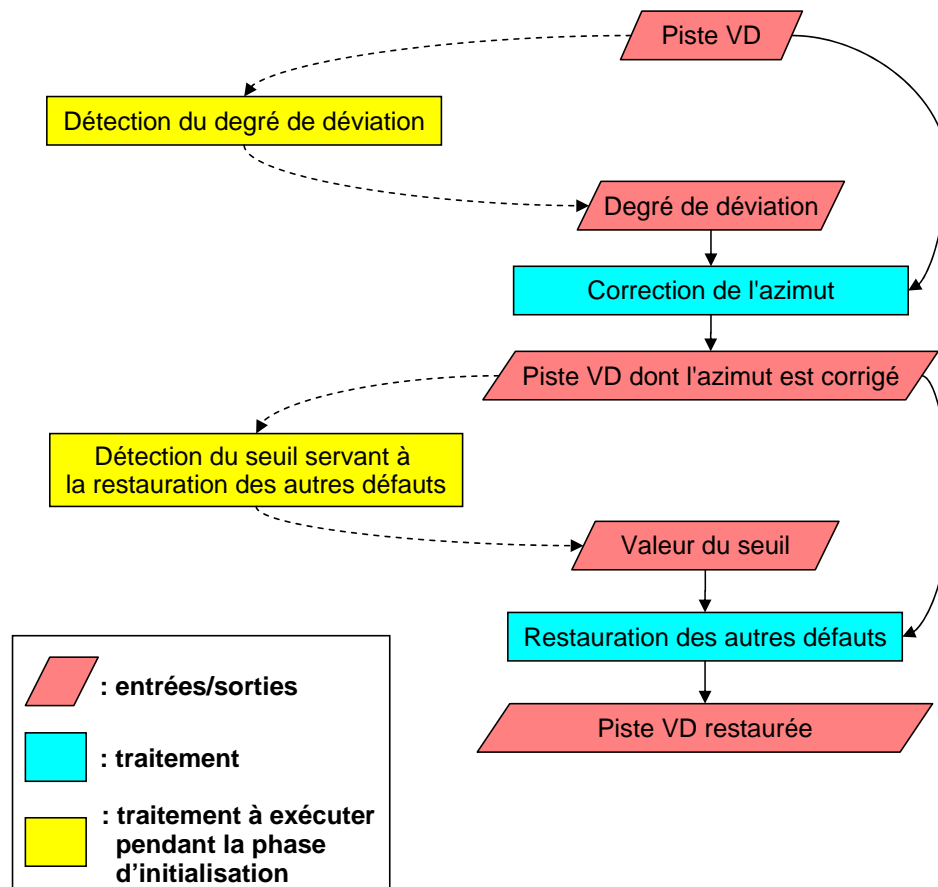


FIGURE 5.3 – Dépendances et entrées/sorties des modules de la restauration des pistes VD.

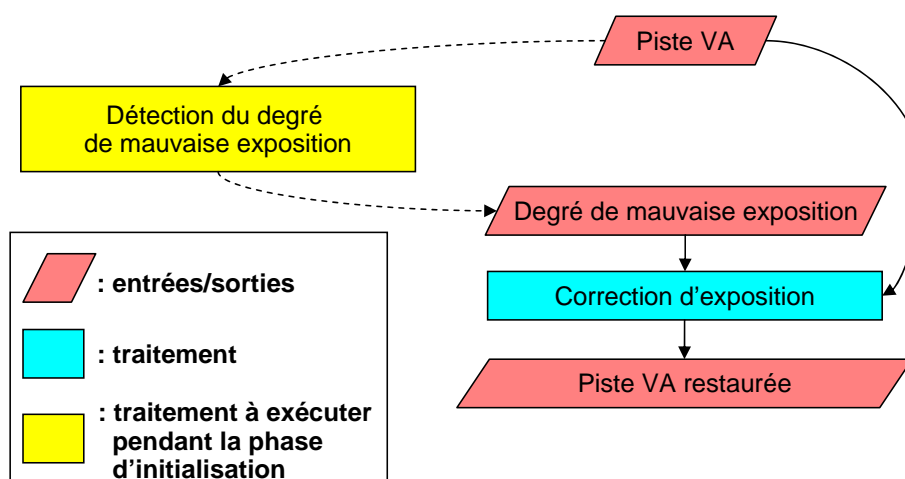


FIGURE 5.4 – Dépendances et entrées/sorties des modules de la correction d'exposition.

Assemblage des deux méthodes de correction d'exposition

Nous avons vu que la correction de la mauvaise exposition pouvait se faire de deux manières différentes. Nous donnons ici plus de détails sur leur intégration dans le système global de restauration.

La méthode basée sur la table de correspondance des niveaux de gris est la méthode la moins gourmande en termes de temps de calcul, mais elle n'est adaptée qu'aux cas des pistes ayant une légère mauvaise exposition. Pour cette raison, nous calculons d'abord la valeur optimale de δ_x (voir la section 4.5.2). Si cette valeur est inférieure ou égale, en valeur absolue, à 0,5, la méthode de correction basée sur la table de correspondance des niveaux de gris est à utiliser. Au delà de cette valeur, cette correction réduit considérablement la dynamique du signal audio final. Pour cette raison, nous calculons dans de tels cas, la taille optimale de l'élément structurant permettant de corriger la mauvaise exposition (voir section 4.5.1). Cet élément structurant sera utilisé pour la restauration de toute la piste son. L'organigramme de la figure 5.5 récapitule toutes ces étapes.

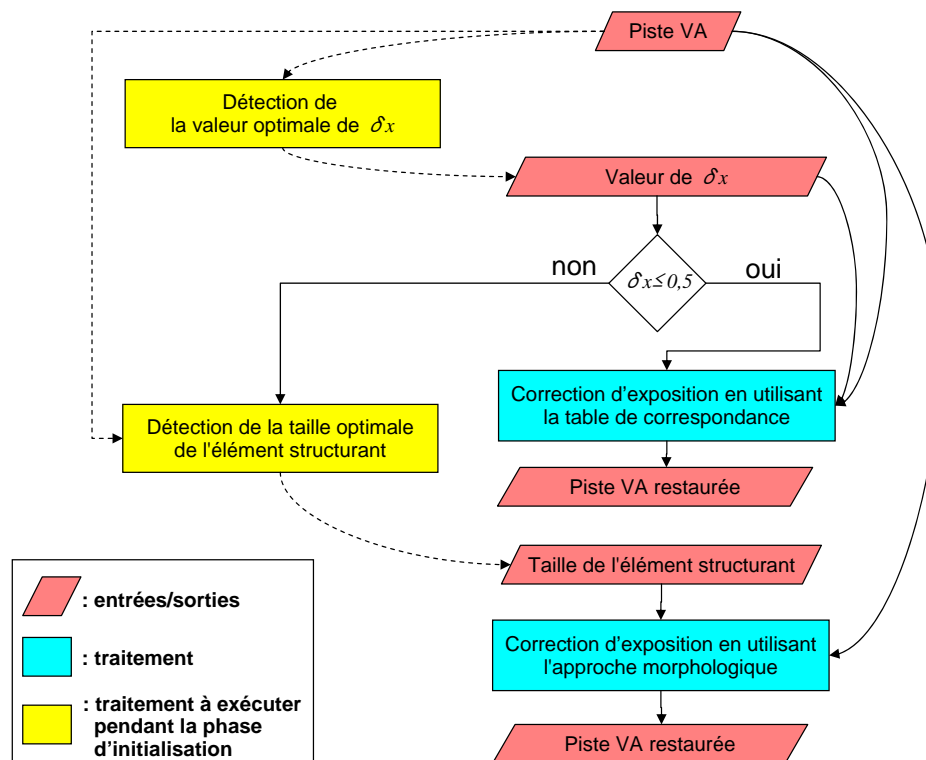


FIGURE 5.5 – Utilisation des deux méthodes de correction d'exposition.

Cas des pistes mal exposées présentant d'autres dégradations

Nous avons traité séparément le problème de la mauvaise exposition et celui des défauts dus au vieillissement du film. Quand ces deux problèmes se présentent en même temps, il est important de restaurer d'abord les défauts dus au vieillissement du film, parce que les indicateurs de détection du degré de mauvaise exposition requièrent une image «propre» et la présence d'éventuels défauts risque de les dérouter considérablement.

Notons cependant que nous n'avons pas été confronté à des pistes son présentant ces deux types de défauts. Nous n'avons, en conséquence, pas pu valider cet assemblage.

5.2 Temps de calcul détaillés

Nous donnons dans cette section les temps de calcul pour chacun des modules développés. Tous ces temps de calcul ont été obtenus en utilisant un processeur Pentium 4 à 2,4 GHz avec 2GB de RAM.

Il est cependant à noter que jusqu'ici, nous ne sommes pas préoccupé de l'optimisation du code.

Restauration des pistes à largeur variable

Dans la phase d'initialisation, la détection du degré de déviation, sur une succession de 10 images, s'effectue en 265 secondes, en considérant toutes les déviations dans $[-10; 10]$ pixels] avec un pas de 0,1 pixel.

Le tableau 5.1 montre les temps de calcul de tous les modules du système de restauration des pistes à largeur variable, pour une longueur de pellicule correspondant à une seconde de projection. Remarquons que la croissance de régions est le module le plus coûteux en termes de temps de calcul.

Restauration des pistes à densité variable

La détection du degré de déviation, sur une succession de 10 images, s'effectue en 55 secondes, en considérant toutes les déviations dans $[-10; 10]$ pixels] avec un pas de 0,1 pixel. La détection de la valeur du seuil s'effectue, quant à elle, en 205 secondes, si l'on considère toutes les valeurs du seuil de 0 à 255.

Les temps de calcul de tous les modules de ce système sont montrés dans le tableau 5.2. Remarquons que ce type de pistes est restauré en temps-réel, puisque la restauration d'une seconde de son se fait en moins d'une seconde.

Module	Temps d'exécution (en secondes)
Détection de l'axe de symétrie	4,0
Correction de l'azimut	0,4
Renforcement de la symétrie	1,0
Croissance de régions	24,3
Renforcement de la causalité	2,1
Lissage des contours	10,5
Total	42,3

TABLE 5.1 – Temps d'exécution du système de restauration des pistes VA sur 75 images, soit une seconde de son.

Module	Temps d'exécution (en secondes)
Correction d'azimut	0,4
Correction des autres défauts	0,2
Total	0,6

TABLE 5.2 – Temps d'exécution du système de restauration des pistes VD sur 75 images.

Restauration des pistes mal exposées

La tableau 5.3 montre les temps d'exécution des différents indicateurs de détection de mauvaise exposition.

Indicateur	Temps d'exécution (en secondes)
<i>InvarianceOuverture</i>	0,2
<i>InvarianceFermeture</i>	0,2
<i>DissymetrieSurfaces</i>	0,4
<i>DissymetrieCourbures</i>	1,0

TABLE 5.3 – Temps d'exécution des indicateurs de détection de mauvaise exposition sur une image 512×640 .

La détection de la valeur optimale de δx prend 75 secondes, si l'on considère toutes les valeurs dans $[0; \pm 5]$ avec un pas de 0,1. La détection de la taille optimale de l'élément structurant prend 57 secondes, si l'on considère toutes les tailles d'éléments structurants dans $[0; 1]$ avec un pas de 0,1 et une résolution dix fois plus importante.

La correction d'exposition en utilisant la valeur optimale de δx prend 0,03 seconde par image, soit deux fois moins rapide que le temps réel.

La correction par approche morphologique est plus lente. Elle s'effectue en deux

étapes : la première étant l'augmentation de résolution, qui prend 4,7 secondes pour une résolution dix fois plus importante, la deuxième étape étant l'érosion ou la dilatation. Sa durée varie en fonction de la taille de l'élément structurant comme le montre le tableau 5.4.

Taille de l'élément structurant rond	Temps d'exécution (en secondes)
0,1	0,1
0,2	0,3
0,3	0,7
0,4	1,3
0,5	1,9
0,6	2,8
0,7	3,7
0,8	4,8
0,9	6,1
1,0	7,4

TABLE 5.4 – Temps d'exécution de l'érosion (ou de la dilatation) en fonction de la taille de l'élément structurant rond, pour une image 512×640 .

Optimisations

Les méthodes développées admettent quelques optimisations.

En ce qui concerne les pistes à largeur variable, la valeur du pli est actuellement calculée pour toutes les lignes, il serait envisageable de la calculer toutes les n lignes, où n est un entier à déterminer empiriquement. Cette optimisation est valable aussi bien pour la détermination de l'axe de symétrie que pour la détection du degré de déviation.

De manière similaire, dans le cas des pistes à densité variable, la variance peut aussi être calculée périodiquement et non pas pour toutes les lignes.

Pour des raisons pratiques, nous avons jusqu'à présent utilisé, pour tous les traitements, la même taille des buffers (640 pixels + éventuellement une zone de superposition). Ceci n'est pas forcément optimal, parce que le fait de traiter plusieurs buffers successifs en tant qu'un seul buffer permet de réduire le nombre de zones de superposition¹, puisque celles-ci seront ajoutées seulement au buffer «global».

1. Rappelons ici qu'à cause des effets de bords, certains traitements nécessitent que chaque image soit complétée par les bords de ses deux images voisines. Nous appelons ces bords ajoutés «zones de superposition»

La réalisation de toutes ces optimisations devrait réduire le temps de calcul global de 20%.

5.3 Implémentation et possibilités de parallélisation

Le système de restauration a été développé en C++ en utilisant *Morph-M*². Bien que la version actuelle du logiciel permette de restaurer une piste son sous forme de séquence d'images enregistrées, certains traitements dont le temps d'exécution n'est pas trop important, peuvent facilement être adaptés pour une restauration en ligne.

Nous avons vu que le système est initialement prévu pour un fonctionnement séquentiel : l'entrée d'un module correspond, le plus souvent, à la sortie d'un autre module. Cette propriété rend la parallélisation du système, une tâche très délicate. Cependant, plusieurs points peuvent être pris en compte pour une éventuelle implémentation sur une architecture multiprocesseur.

Traitements locaux et traitement globaux

Plusieurs traitements correspondent à des «opérateurs locaux». Rappelons qu'un opérateur est dit local, si le résultat de son application à une partie de l'image est le même, que cette partie soit traitée séparément ou dans l'image entière. Si ce résultat est différent, l'opérateur est alors dit «global».

L'implémentation d'un opérateur local est plus facilement réalisable sur une architecture multiprocesseur qu'un opérateur global. Il suffit en effet de diviser l'image d'entrée de l'opérateur local en parties égales sur tous les processeurs.

Notons que certains opérateurs locaux, comme l'érosion ou la dilatation par exemple, ne traitent pas les pixels de l'image indépendamment, mais en fonction de leurs voisinages. Dans le cas de tels opérateurs, le résultat du traitement est le même que l'image soit traitée séparément ou globalement, sauf sur les bords de la séparation, à cause des effets de bord. Ce type de traitements sera appelé ci-après «opérateur quasi-local», il se parallélise de la même manière que les opérateurs locaux, à l'exception qu'il requiert des zones de superposition. La taille de ces zones de superposition change en fonction du traitement, comme expliqué précédemment.

Le seul traitement global dans notre chaîne de restauration est la croissance de régions. Nous avons vu précédemment qu'en pratique, la croissance de régions «converge» vers un opérateur quasi-local, à partir d'une zone de superposition de taille 30 (voir section 2.3.4).

2. <http://morphm.ensmp.fr/>, Morph-M est le fruit de travail de plusieurs chercheurs au Centre de Morphologie Mathématique. Cette librairie offre des outils puissants de morphologie mathématique et de traitement d'images.

Cette remarque est valable pour des images d'une longueur de 640 pixels. Pour des images ayant une longueur plus petite, la taille «nécessaire» de la zone de superposition risque d'être importante par rapport à la taille de l'image elle-même, ce qui rend inefficace la parallélisation de la croissance de régions par la même méthode que celle des opérateurs quasi-locaux.

Il n'existe pas de méthode générale de parallélisation des opérateurs globaux, la méthode de parallélisation change en fonction du traitement. En ce qui concerne la croissance de régions, plusieurs algorithmes parallèles ont été proposés ([58, 5]). Dans ces méthodes, la propagation de chaque région est effectuée par un processeur et la synchronisation est assurée par des algorithmes de contrôle de vitesse de propagation et de gestion des collisions.

Temps d'exécution de chaque traitement

Les temps d'exécution ne sont pas du même ordre pour tous les traitements. La croissance de régions, par exemple, est soixante fois plus lente que la correction de l'azimut. Dans un système séquentiel, la lenteur d'un traitement paralyse toute la chaîne. Pour cette raison, il est important de donner la priorité aux traitements les plus lents lors de la parallélisation du système.

5.4 Conclusion

Dans ce chapitre, nous avons rappelé tous les modules développés pour la restauration des pistes à largeur et à densité variable ainsi que pour la restauration des pistes mal exposées.

Nous avons détaillé les entrées/sorties de chacun de ces modules et montré toutes les dépendances qui existent entre eux. Nous avons notamment distingué les modules servant à la paramétrisation du système de ceux qui servent à la restauration proprement dite.

Nous avons aussi donné les temps de calcul détaillés et mis en évidence les étapes les plus lentes. Enfin, nous avons donné plusieurs recommandations quant à l'implémentation du système de restauration sur une architecture multiprocesseur.

Tous les modules développés sont en cours d'optimisation et d'intégration pour permettre une restauration en temps-réel. Une restauration hors-ligne est aussi prévue, notamment pour les pistes à largeur variable présentant de sévères dégradations, ou un important degré de mauvaise exposition.

Conclusion générale

6.1 Apports de cette thèse

La restauration de la piste sonore optique est généralement effectuée dans le domaine audio. Dans cette thèse, nous avons proposé plusieurs solutions originales pour effectuer cette restauration directement dans le domaine image. Les méthodes que nous avons développées permettent de restaurer les pistes à largeur variable et les pistes à densité variable. Nous avons, non seulement pris en compte les dégradations dues au vieillissement du film, mais aussi celles qui sont dues au problème de mauvaise exposition.

Pour la restauration des pistes à largeur variable, nous avons développé une méthode basée sur le pli de l'image pour détecter avec précision l'axe de symétrie. Nous avons également proposé un algorithme précis pour la correction de l'azimut. Nous avons assuré le nettoyage de l'image dans une étape de segmentation par croissance de régions. Nous avons aussi considéré d'autres étapes pour la correction de la géométrie de la piste son. Enfin, dans une étape d'anti-aliasing, nous avons utilisé l'image initiale pour introduire plus de dynamique dans l'image segmentée [31].

En ce qui concerne la restauration des pistes à densité variable, nous avons également développé un algorithme de correction d'azimut. Nous avons éliminé les défauts par une méthode de seuillage efficace, dans laquelle la valeur du seuil minimise la variation totale du signal audio résultant, ce qui nous a permis de réduire au maximum le bruit [32].

Pour détecter la mauvaise exposition, nous avons montré que ce problème est très proche d'une érosion (ou d'une dilatation) morphologique, et peut en conséquence, être détecté en vérifiant l'invariance par rapport à l'ouverture (ou la fermeture) morphologique. Nous avons également montré que la mauvaise exposition introduit des dissymétries entre les pics et les vallées de la piste son, et que cette dissymétrie peut être mesurée au niveau des surfaces ou au niveau des courbures.

Nous avons montré que la dilatation (ou l'érosion) morphologique constitue une première restauration de la mauvaise exposition ([7, 30]). Pour avoir une restauration précise, nous avons montré comment effectuer ces opérations morphologiques avec un élément structurant de taille réelle, et comment choisir cette taille d'élément structurant en fonction du niveau détecté de mauvaise exposition. Nous avons également proposé une autre méthode de correction, pour les pistes légèrement dégradées, en appliquant une table de

correspondance de niveaux de gris selon le niveau de mauvaise exposition détecté.

Nous avons aussi abordé le problème de la communication entre les différents modules développés ainsi que leur parallélisation et leur intégration dans un système global de restauration.

Les résultats de toutes ces méthodes de restauration ont été validés par les archivistes du film d'un côté, et par une campagne de tests d'écoute en aveugle d'un autre côté.

6.2 Perspectives

Les tests d'écoute réalisés nous ont permis d'avoir une première évaluation de nos méthodes de restauration. D'autres tests plus complets qui répondent aux recommandations de l'Union Internationale de Télécommunication méritent d'être conduits pour avoir une évaluation plus précise, notamment en ce qui concerne les comparaisons dans lesquelles les différences sont difficilement décelables.

La restauration des pistes à largeur variable reste très lente par rapport à celle des pistes à densité variable. Nous pensons que des techniques de restauration plus rapides doivent être mises en place, notamment pour la restauration des pistes légèrement dégradées. Dans ce sens, un brevet sur des techniques de restauration temps-réel est actuellement en cours de rédaction à l'Université de la Rochelle.

La restauration des pistes à densité variable a été validée sur une piste simulée ; sa robustesse par rapport à des pistes très dégradées doit être étudiée.

Même si la plupart des paramètres du système ont été maintenus constants pour toutes les pistes son sur lesquelles nous avons travaillé, nous pensons qu'un travail supplémentaire sur la paramétrisation automatique de la méthode simplifierait davantage la tâche de l'utilisateur final.

Ces travaux ont suscité l'intérêt de la presse spécialisée. Un article a été récemment publié dans la revue SONOVISION sur les différents aboutissements du projet RESONANCES [46].

Les Archives Françaises du Film ainsi que Général des Travaux Cinématographiques (GTC) seront bientôt équipés de versions «clonées» du défileur RESONANCES ainsi que du système de restauration développé.

Bibliographie

- [1] R. Adams and L. Bischof. Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.*, pages 641–647, 1994. [31](#)
- [2] L. Alfonsi. *Le cinéma du futur : les enjeux des nouvelles technologies de l'image*. Les Presses de l'Université Laval, 2005. [2](#)
- [3] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel. Axiomatisation et nouveaux opérateurs de la morphologie mathématique. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 1992. [74](#)
- [4] ATMEL. Flat field correction, how flat field correction and contrast expansion are implemented in aviiva cameras. ATMEL Application Note, 2005. [42](#)
- [5] D. A. Bader, J. JaJa, D. Harwood, and L. S. Davis. Parallel algorithms for image enhancement and segmentation by region growing with an experimental study. In *10th International Parallel Processing Symposium (IPPS '96)*, 1996. [94](#)
- [6] T. Balio. *The American Film Industry*. University of Wisconsin Press, 1985. [2](#)
- [7] B. Besserer, A. Hassaïne, and E. Decencière. Restoration of under/overexposed optical soundtracks. In *Archiving*, pages 83–88, Bern - Switzerland, June 2008. [95](#)
- [8] S. Beucher. Numerical residues. In *Proceedings of the 7th International Symposium on Mathematical Morphology*, Paris, 2005. [18](#)
- [9] S.A. Bhukhanwala and T.V. Ramabadran. Automated global enhancement of digitized photographs. *IEEE Transactions on Consumer Electronics*, 40(1) :1–10, 1994. [56](#), [57](#), [79](#)
- [10] H. Blum. A transformation for extracting new descriptors of shape export. In *Models for the Perception of Speech and Visual Form*, 1967. [74](#)
- [11] G. Borgefors. Distance transformations in digital images. *Computer Vision, Graphics and Image processing*, 34(3) :344–371, 1986. [32](#)
- [12] G. Borgefors. *On hierarchical edge matching in digital images using distance transformations*. PhD thesis, Royal Institute of Technology, Stockholm, 1986. [32](#)
- [13] J. E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1) :25–30, 1965. [27](#)
- [14] R. W. Brockett and P. Maragos. Evolution equations for continuous-scale morphological filtering. *IEEE transactions on signal processing*, 1994. [74](#)
- [15] E. Brun. Restauration par traitement d'images de la piste son optique des films cinématographiques. Rapport de stage de fin d'études. Centre de morphologie mathématique, 2006. [18](#)

- [16] E. Brun, A. Hassaïne, B. Besserer, and E. Decenci re. Restoration of variable area soundtracks. In *Proceedings of the 2007 IEEE International Conference on Image Processing*, San Antonio, Texas, USA, 2007. 18, 30
- [17] H. Carlsen and D. George. Apparatus for reading analog sound tracks on film with circuit for averaging sound track boundaries to reduce the effect of noise. United States Patent 5526075, 1996. 16
- [18] S. Chen and R.M. Haralick. An automatic algorithm for text skew estimation in document images using recursive morphological transforms. In *Proceedings of the 1994 IEEE International Conference on Image Processing*, 1994. 23
- [19] B. Coe. *The history of movie photography*. Eastview Editions, 1981. 3
- [20] J. F. Corso. Age and sex differences in pure-tone thresholds. *Arch Otolaryngol*, 4(77) :385–405, 1961. Survey of Hearing Levels From 18 to 65 Years. 10
- [21] D. Crafton. *The talkies : American cinema’s transition to sound, 1926-1931*. Charles Scribner’s Sons, 1997. 2
- [22] R. Davis and F. M. Walters. Sensitometry of photographic emulsions and a survey of the characteristics of plates and films of american manufacture. *Scientific Papers of the Bureau of Standards*, 18(439), 1922. 53
- [23] E. Decenci re. *Restauration automatique de films anciens*. PhD thesis, Ecole Nationale Sup rieure des Mines de Paris, 1997. 1
- [24] S. Derrode and F. Ghorbel. Shape analysis and symmetry detection in gray-level objects using the analytical fourier-mellin representation. *Signal Processing*, 84(3) :25–39, 2004. 20
- [25] Dolby. The evolution of dolby film sound. Dolby Technical report, 2005. 2
- [26] R. Enficiaud. *Algorithmes multidimensionnels et multispectraux en morphologie math matique : Approche par m ta-programmation*. PhD thesis, Ecole des Mines de Paris, 2007. 18, 63
- [27] J. G. Frayne. *Elements of sound recording*. John Wiley and Sons, Inc., 1949. 3, 23, 55
- [28] S. J. Godsill and P. J. W. Rayner. *Digital Audio Restoration - A Statistical Model-Based Approach*. Springer-Verlag, 1998. 6
- [29] F. Guichard, P. Maragos, and J.-M. Morel. *Space, Structure and Randomness*, volume 183 of *Lecture Notes in Statistics*, chapter Partial Differential Equations for Morphological Operators, pages 369–390. Springer Verlag, 2005. Contributions in Honor of Georges Matheron in the Fields of Geostatistics, Random Sets, and Mathematical Morphology. 74
- [30] A. Hassaïne, B. Besserer, and E. Decenci re. Etude des pistes son optiques sur/sous-expos es. *Soumis   la revue «Traitement du signal»*, 2009. 95
- [31] A. Hassaïne, E. Decenci re, and B. Besserer. Efficient restoration of variable area soundtracks. *Image Analysis and Stereology*, pages 113–119, June 2009. 95

- [32] A. Hassaïne, E. Decencière, and B. Besserer. Restoration of variable density film soundtracks. In *Proceedings of the 17th European Signal Processing Conference*, 2009. 95
- [33] J. Hull. Surround sound past, present, and future. Technical report, Dolby Laboratories Inc., 2005. A history of multichannel audio from mag stripe to Dolby Digital. 2
- [34] ISO. A-chain frequency response for reproduction of 35 mm photographic sound – reproduction characteristics. International standard, 1986. 4, 8
- [35] ISO. B-chain electro-acoustic response of motion-picture control rooms and indoor theatres – specifications and measurements. International standard, 1987. 4
- [36] ISO. Acoustics – statistical distribution of hearing thresholds as a function of age. International standard, 2000. 10
- [37] R. C. Jones. On the point and line spread functions of photographic images. *Journal of optical society of America*, 48(12), 1958. 56
- [38] M. F. Kelly and M. D. Levine. Annular symmetry operators : A method for locating and describing objects. In *Proceedings of the IEEE International Conference on Computer Vision*, 1995. 20
- [39] N. Kiryati and Y. Gofman. Detecting symmetry in gray level images : The global optimization approach. *International Journal of Computer Vision*, 29 :29–45, 1998. 20
- [40] Kodak. Cross-modulation distortion testing for the motion picture laboratory. Eastman Kodak Technical report, 2001. 54, 55
- [41] A. C. Kokaram. *Motion Picture Restoration : Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*. Springer-Verlag, 1998. 1
- [42] A. Kuiper. Detection of dirt blotches on optical soundtracks using digital image processing. In *15th International Czech - Slovak Scientific Conference Radioelektronika*, 2005. 18
- [43] A. Kuiper. Removal of blotches on multiple double sided variable area code sound tracks using matlab. In *Proceedings of conference Student EEICT 2006*, Czech Republic, 2006. 18
- [44] A. Lovichi. *Sensitométrie des films sonores*. KADAK-PATHÉ VINCENNES, 1943. 53
- [45] G. Messina, A. Castorina, S. Battiato, and A. Bosco. Image quality improvement by adaptive exposure correction techniques. In *International Conference on Multimedia and Expo*, 2003. 57
- [46] L. Ollier. Resonances restaure le son à sa source. *SONOVISION DIGITAL FILM*, pages 19–21, mai 2009. 96

- [47] T. Oohashi, N. Kawai, E. Nishina, M. Honda, R. Yagi, S. Nakamura, M. Morimoto, T. Maekawa, Y. Yonekura, and H. Shibasaki. The role of biological system other than auditory air-conduction in the emergence of the hypersonic effect. *Brain Research*, 2006. 10
- [48] T. Oohashi, E. Nishina, M. Honda, Y. Yonekura, Y. Fuwamoto, N. Kawai, T. Maekawa, S. Nakamura, H. Fukuyama, and H. Shibasaki. Inaudible high-frequency sounds affect brain activity : Hypersonic effect. *Journal of Neurophysiology*, 2000. 10
- [49] S. Osher and L. I. Rudin. Feature-oriented image enhancement using shock filters. *IEEE transactions on image processing*, 27(4) :919–940, 1990. 47
- [50] D. Poetsch, D. Richter, and J. Huegel. Device and method for reproducing photographic sound recordings. United States Patent Application 20030165087, 2003. 16
- [51] V. Prasad and B. Yegnanarayana. Finding axes of symmetry from potential fields. *IEEE Trans. PAMI*, 13(12) :1559–1566, 2004. 20
- [52] P. Read and M.-P. Meyer. *Restoration of Motion Picture Film*. Butterworth-Heinemann, 2000. 1, 5
- [53] D. Richter, I. Kurreck, and D. Poetsch. Restoration of optical variable density sound track on motion picture films by digital image processing. In *Proceedings of the 7th International Conference on Optimization of Electrical and Electronic Equipments*, 2002. 42
- [54] D. Richter, D. Poetsch, and A. Kuiper. Localization of faults in multiple double sided variable area code sound tracks on motion picture films using digital image processing. In *13th International Czech - Slovak Scientific Conference Radioelektronika*, 2003. 17
- [55] J. Serra. *Image Analysis and Mathematical Morphology*, volume 1. Academic Press, London, UK, 1982. 57
- [56] J. A. Sethian. *Level Set Methods : Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, 1996. 74
- [57] D. Shen, H. Ip, K. Cheung, and E. Teoh. Symmetry detection by generalized complex (gc) moments : A close-form solution. *IEEE Trans. PAMI*, 21(5) :466–476, 1999. 20
- [58] D.E. Singh, D.B. Heras, and F.F.Rivera. Parallel seeded region growing algorithm. In *Proc.VIII Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes.SNRFAI'99*, 1999. 94
- [59] P. Streule. Digital image based restoration of optical movie soundtracks. Master's thesis, ETH Zurich, 1999. 17, 23, 24, 43
- [60] M. Su and C. Chou. A modified version of the k-means algorithm with a distance based on cluster symmetry. *IEEE Trans. PAMI*, 23(6), 2001. 20
- [61] C. Sun. Symmetry detection using gradient information. *Pattern Recognition Letters*, 16(9) :987–996, 1995. 20

- [62] S. Takeda, I. Morioka, K. Miyashita, A. Okumura, Y. Yoshida, and K. Matsumoto. Age variation in the upper limit of hearing. *European Journal of Applied Physiology*, 65(5) :403–408, 1992. 10
- [63] J. Taquet, B. Besserer, A. Hassaïne, and E. Decenci re. Detection and correction of under/overexposed optical soundtracks by coupling image and audio signal processing. *EURASIP Journal on Advances in Signal Processing*, 2008. 58
- [64] H. M. Tremaine. *The Audio Cyclopedia*. Howard W. Sams and Co., 1959. 23
- [65] UIT. M thodes d  valuation subjective des d gradations faibles dans les syst mes audio y compris les syst mes sonores multivoies. Technical report, Union internationale des t l communications, 1997. 12
- [66] J. A. Valenzuela. Digital reproduction of variable density film soundtracks. United States Patent Application 20060232745, 2006. 42
- [67] J. A. Valenzuela and V. R. Williams. Processing variable area film soundtracks. United States Patent Application 20040189946, 2004. 16
- [68] J. A. Valenzuela and V. R. Williams. Digital reproduction of optical film soundtracks. United States Patent 7050142, 2006. 16
- [69] J. A. Valenzuela and V. R. Williams. Variable area film soundtrack renovation. United States Patent 7126663, 2006. 56
- [70] J. A. Valenzuela and V. R. Williams. Digital imaging of optical film soundtracks. United States patent 7403258, 2008. 16
- [71] R. van den Boomgaard. *Mathematical Morphology : Extensions towards Computer Vision*. PhD thesis, University of Amsterdam, 1992. 74
- [72] S.V. Vaseghi. *Advanced Signal Processing and Noise Reduction*. John Wiley and Sons, Inc., 2000. 6
- [73] H. Wang and D. Suter. Ltsd : A highly efficient symmetry-based robust estimator. In *International Conference on Control, Automation, Robotics and Vision*, volume 1, pages 332–337, 2002. 20
- [74] T. Yuan and X. Tang. Efficient local reflectional symmetries detection. In *Proceedings of the 2005 IEEE International Conference on Image Processing*, 2005. 20
- [75] H. Zabrodsky, S. Peleg, and D. Avnir. A measure of symmetry based on shape similarity. In *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR ’92., 1992 IEEE Computer Society Conference on*, pages 703–706, June 1992. 20